Are gridded precipitation datasets a good option for streamflow simulation across the Juruá river basin, Amazon?

Frédéric Satgé, Benjamin Pillot, Henrique Roig, Marie-Paule Bonnet

Version of Record: https://www.sciencedirect.com/science/article/pii/S0022169421008234
Manuscript_baca05537df0e223597ae162b991597a

Are gridded precipitation datasets a good option for streamflow simulation across the Juruá river basin, Amazon?

Frédéric Satgé[1*], Benjamin Pillot[1], Henrique Roig[2] and Marie-Paule Bonnet[1]

[1] ESPACE-DEV, Univ Montpellier, IRD, Univ Antilles, Univ Guyane, Univ Réunion, Montpellier, France

[2] Instituto de Geociência (IG), Universidade de Brasília, Brasilia-DF, Brazil

*Corresponding author: frederic.satge@ird.fr

Abstract

Over the past few decades, many gridded P-datasets with nearly global coverage have become available. These P-datasets offer an unprecedented opportunity to constrain hydrological modeling in remote region where the gauge network is sparse. However, few studies report on P-datasets reliability for discharge simulations limiting the use of these datasets. This study investigates the reliability of available gridded P-datasets for streamflow simulations for 10 basins of the Juruá watershed, located in the Amazon region. A total of 19 P-datasets including both satellite-based and reanalysis-based precipitation estimates, are considered to provide a comprehensive overview of currently available options. Used as forcing data in two lumped hydrological models (GR4j and HyMOD), some P-datasets led to a more realistic simulation of daily and monthly streamflow than the simulation based on precipitation estimates derived from the gauges network. P-dataset ranking depends on the considered basin and time step (i.e. daily, monthly), suggesting variability in spatial reliability for all considered P-datasets. In addition, the P-dataset reliability increases with the surface area of the considered basins. This can be partially explained by the aggregation of precipitation on larger spatial scales counterbalancing potential spatial inconsistencies at more local scales in the P-dataset, and by the better modeling of smoother hydrographs at the outlets of larger basins. Overall, IMERG-F v.6 and CMORPH-BLD appear to be the most efficient P-datasets for the region under consideration.

1. Introduction

Global warming and increasing anthropogenic pressure threaten the sustainability of water resources in many parts of the world and undermine the integrity of the ecosystems and societies on which they depend. In this context, observing and quantifying precipitation and its evolution over time is particularly important for estimating the recharge processes of surface and groundwater systems with respect to their use. Precipitation is traditionally estimated from networks of in-situ stations whose data are still collected manually in many regions, then digitized, and therefore exposed to reporting errors. The spatial distribution of the stations, constrained by the maintenance costs depending on the accessibility to the installation sites, is very heterogeneous, and leads to vast areas without any observation. The high maintenance cost is often prohibitive for many countries. This results in substantial gaps in the surveyed time series and to sparsely populated station networks usually distributed among easily accessible areas. The precipitation estimate in remote ungauged areas is then based on the interpolation of scattered and remote stations, which are not very representative of the local precipitation dynamics. Similar difficulties are common in transboundary regions where international conflicts on water resources may limit the access to the data from the national monitoring networks. Thus, in this context, open-access gridded precipitation datasets (P-datasets) providing near-global spatial coverage stand for a particularly attractive alternative.

Over the past 20 years, 30 near global-scale P-datasets with different space–time coverage and resolution have been identified (Sun et al., 2018). On a general way, the precipitation estimates are derived from different inputs such as gauges observation (i.e. CPC, CRU, GPCC) and/or satellite information (i.e. IMERG, TMPA, CMORPH) and/or physical and dynamical models output (reanalysis dataset) (i.e. ERA5, MERRA2, WFDEI). Therefore, P-datasets estimates present discrepancies in space and time depending on the inputs sensitivity to local climate and topographic context. For instance, gauge-based precipitation estimates reliability is closely related to the gauges network density and distribution used for the interpolation process (Sun et al., 2015). Satellite-based precipitation estimates rely on PMW and IR sensors which measurements are not only influenced by precipitations but also by quick change in topographic and surface emissivity (Ferraro et al., 1998; Levizzani et al., 2002). Satellite-based precipitation estimates are generally less reliable across mountainous region (Hussain et al., 2017; Satgé et al., 2017), snow covered area and big lake/river regions (Paiva et al., 2011; Satgé et al., 2016; Tian and Peters-Lidard, 2007). Moreover, the irregular sampling of the

satellites hardly capture short-term and slight precipitation events (Gebregiorgis and Hossain, 2013; Tian et al., 2009). Finally, reanalysis-based precipitation estimates reliability vary in space and time as the models are generally more adapted for large-scale stratiform systems than small-scale convective precipitation cells (i.e. Beck et al., 2019, Satgé et al., 2020).

In this context, P-dataset estimates are generally compared to gauges observations to report on P-dataset space and time reliability across different regions (i.e. Beck et al., 2019; Maggioni et al., 2016; Maggioni and Massari, 2018; Satgé et al., 2020). In gauge-scarce regions, this comparison is limited as only one gauge per grid-cell is usually available for comparison. As a result, many precipitation events observed at the grid-cell level (areal measurement of P-datasets) may be lost or underestimated at the gauge level (point measurement of gauges) introducing uncertainties in the comparison (i.e. Salles et al., 2019; Satgé et al., 2019; Tang et al., 2018). To overcome this issue, an alternative approach is to compare the simulated streamflow (using P-datasets as forcing data) with the observed streamflow. Because streamflow is function of basin-wide precipitation estimates, this comparison overcomes the spatial scale discrepancy between grid-cell and rain gauges observation. Based on this method, recent studies use P-datasets as forcing data for streamflow modelling to assess the reliability and suitability of P-datasets for streamflow modelling.

Accordingly, Satgé et al. (2019) assessed 7 satellite-based P-datasets (SM2RAIN-CCI, SM2RAIN-ASCAT, IMERG-E, -L and –F v.6, CHIRPS v.2 and MSWEP v.2.2) over 2 basins located in the south American Andean plateau. Results show that daily streamflow modelled with MSWEP v.2.2 and IMERG-F were in close agreement with the observed one. Another study compared 7 P-datasets including satellite-based (IMERG-E, IMERG-F, PRISM-SMAP, PRISM-SMOS, GPM+SM2Rain), reanalysis-based (ERA5) and gauge-based (GPCC) ones for streamflow modelling of 10 basins across Europe and Africa (Brocca et al., 2020). This work shows that P-datasets offer a great opportunity for streamflow modelling over scarcely gauged basins. Across 15 basins located across the European Mediterranean Sea region, 4 satellite-based P-datasets (TMPA-RT v.7, CMORPH, PERSIANN, SM2RAIN-CCI) provided less realistic streamflow simulation than precipitation gauges observations (Camici et al., 2018). Additionally, 3 P-datasets (TMPA-RT, CMORPH and SM2RAIN-ASCAT) were evaluated for the streamflow modelling of 1318 European basins (Camici et al., 2020). According to the authors, these P-datasets can reasonably be used as forcing data for the considered basins. Jiang and Bauer-Gottwein, (2019) assessed 3 P-datasets (TMPA-adj v.7, IMERG-E and –L v.6) for streamflow modelling of 300 Chinese basins and

found that IMERG P-datasets provide comparable streamflow simulations to gauge-based precipitation estimates. Tarek et al., (2020) used ERA5 P-dataset to force hydrological modelling for 3138 basins located across North America. Results show that ERA5-based hydrological modelling performance is equivalent to the one using gauges observations over most of the considered basins.

In the above-described context, P-datasets represent a clear opportunity for streamflow modelling especially across remote regions where gauge information is often lacking or missing. However, the current state of the art only considers a limited number of P-datasets (up to 7 in the above example). With 30 P-datasets recently listed (Sun et al., 2018) there is a need for considering a more exhaustive P-dataset sample in order to provide a comprehensive reliability overview of the available P-datasets. In this line of work, 17 and 18 P-datasets including satellite, reanalysis and gauges-based P-datasets were considered for hydrological modelling across the West African basin of Volta basin (Dembele et al., 2020) and 8 large scale basins (Mazzoleni et al., 2019), respectively. According to Mazzoleni et al. (2019), there is no unique best performing P-dataset for all basins and results strongly depend on the basin characteristics.

In this context, the present study assesses 19 P-datasets for hydrological modelling across 10 basins located in the Juruá watershed, one of the large tributary streams of the Amazon River. The analysis is conducted at both daily and monthly time steps and aims at providing important feedback to support P-dataset selection by potential users, and improve development of the next-generation P-datasets.

2. Materials

2.1. Study Area

Juruá basin covers 188,290 km^2 of which 93.8 % is in Brazil and 6.2 % in Peru. The basin has a mean elevation of 203 m ranging from 18 to 760 m in the Andean region forming the western border with Peru. The mean annual basin precipitation is 2,260 mm·year^{-1} with a rainy season from October to April and a dry season from May to September contributing to 77.5 % and 22.5 % of total annual precipitations.

Ten basins of the Juruá river watershed are considered. They correspond to the areas drained by each streamflow gauge available in the Juruá watershed. Their surface area ranges from 3,439 km^2 to 164,611 km^2 at the streamflow station *Seringal bom futuro* and *Gavião*, respectively (Figure 1c).

2.2. Reference observation

We used the daily gridded precipitation datasets for Brazil firstly developed by Xavier et al. (2015) and updated to version 2.1 by Xavier et al., (2017). This dataset provides precipitation on a regular 0.25° grid for the 1980-2015 period. Precipitation estimates are derived from the interpolation (angular distance weighting) of 9,259 rain gauges operated by the *Agência Nacional de Aguas* (ANA) and the *Instituto Nacional de Meterologia* (INMET), from which 36 are included in the Juruá basin.

The Catchment Attributes and MEteorology for Large-sample Studies datasets for Brazil (CAMELS-BR) is used for reference streamflow observations (Chagas et al., 2020). CAMELS-BR consists in 3,679 streamflow gauges with daily observations operated by the ANA from which 25 are located in the Juruá basin. Each series comes with a daily flag indicating whether the observation was quality checked by the ANA. We used this flag to select gauges with more than 95% of daily observations for the 2001-2010 period for a total of 10 streamflow gauges (Figure 1).

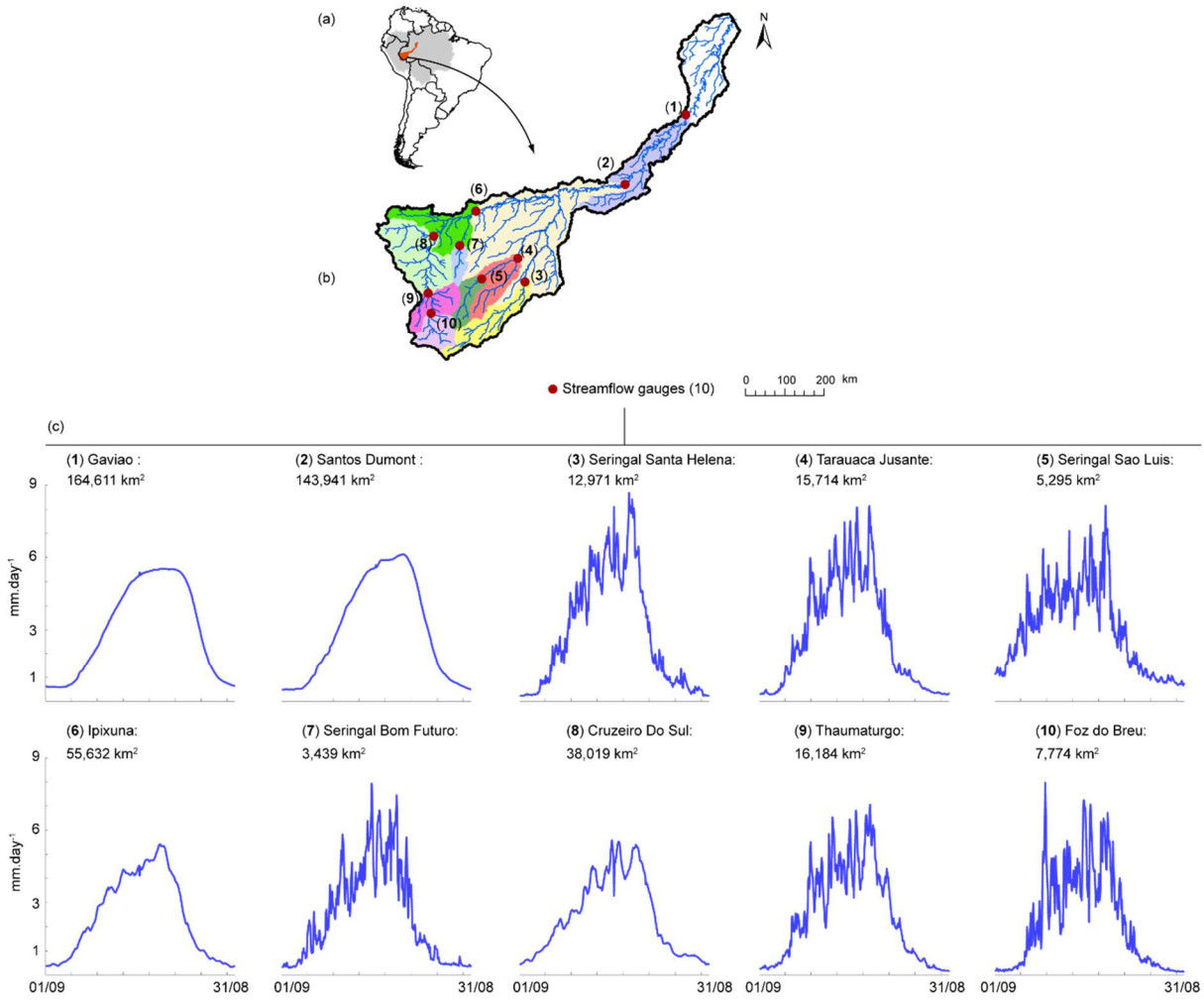


Figure 1. Juruá basin location into the Amazon system (a) with considered sub-basins location (b) and corresponding mean daily streamflow's record over the 2001-2010 period (c). Also reported the drained area (km²) at each considered station

2.3. Gridded Precipitation datasets

This study evaluates 19 P-datasets covering the 10 years period (2001–2010) with streamflow data for the 10 basins. P-datasets are based on three main types of input data to retrieve precipitation:

- Satellite-based precipitation estimates derived from passive/active microwaves and infrared sensors
- Reanalysis-based precipitation estimates derived from physical and dynamical models
- Gauge-based precipitation datasets (i.e. CPC, CRU, GPCC, GPCP, CHPclim and WorldClim2)

Gauge-based precipitation datasets are generally used to adjust the precipitation estimates derived from satellite-based and/or reanalysis-based precipitation estimates. Therefore, the P-datasets considering gauge-based precipitation datasets are expected to provide more reliable precipitation estimates than the other P-datasets. In this context, the 19 P-datasets under consideration in this study were divided into 3 main groups:

- P-datasets relying only on satellite-based and/or reanalysis-based estimates (CHIRP v.2, ERA5, GSMaP-RT v.6, IMERG-E v.6, IMERG-L v.6, MERRA2-FLX and TMPA-RT v.7),
- P-datasets considering gauge-based precipitation datasets and only satellite-based or reanalysis-based (CMORPH-BLD, CMORPH-CRT, GSMaP-Adj v.6, IMERG-F v.6, MERRA2-LND, PERSIANN-CSS-CDR, PERSIANN-CDR, TMPA-Adj v.7, WFDEI-CRU, WFDEI-GPCC)
- P-datasets including satellite-based, reanalysis-based and gauge-based precipitation estimates (CHIRPS v.2 and MSWEP v.2.2)

Finally, it is worth mentioning that P-dataset estimates are available with temporal latency ranging from hours to months after the observation. This is an important feature to take into consideration as P-datasets with short temporal latency are adapted for near-real time data requirement (e.g. flood or landslides forecasting, water resource management for agriculture) while the others are better suited to retrospective climate studies. Also, P-datasets without gauge-based information are available with the shorter temporal latency which ranges from few hours (IMERG-E and -L v.6) to few days (CHIRP v.2, TMPA-RT v.7, GSMaP-RT v.6). With few hours' temporal latency, IMERG-E and -L v.6 would be particularly adapted for near-real time requirement. For more information about the considered P-datasets, readers should refer to the main references listed in Table 1.

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.

Acronym	Full Name	Data	TemporalCoverage	Spatial Resolution	Temporal latency	References
CHIRP v.2	Climate Hazards Group InfraRed v.2	S, R	1981-Present	0.05°	2 days	(Chris Funk e
CHIRPS v.2	CHIRP with stations v.2	S, R, G	1981-Present	0.05°	1 month	(Chris Funk e
CMORPH-CRT	Climate Prediction Center MORPHing bias corrected	S, G	1998-Present	0.25°	6 months	(Xie et al., 20
CMORPH-BLD	CMORPH satellite-gauge merged	S, G	1998-Present	0.25°	1 month	(Xie et al., 20
ERA5	European Centre for MediumRange Weather Forecasts fith generation	R	1981-Present	0.1°	1 month	Muñoz Sabat
GSMaP-RT v.6	Global Satellite Mapping of Precipitation standard v.6	S	2000-Present	0.1°	3 days	(Ushio et al., Yamamoto a 2014)

GSMaP-Adj v.6	GSMaP Adjusted v.6	S, G	2000-Present	0.1°	3 days	(Ushio et al., Yamamoto a 2014)
IMERG-E v.6	Integrated Multi-Satellite Retrievals for GPM Early Run v.6	S	2000-Present	0.1°	4 hours	(Huffman et a
IMERG-L v.6	IMERG-Late Run v.6	S	2000-Present	0.1°	12 hours	(Huffman et a
IMERG-F v.6	IMERG-Final Run v.6	S, G	2000-Present	0.1°	3 months	(Huffman et a
MERRA2-FLX	Modern-Era Retrospective Analysis for Research and Applications 2	~~S~~, R, ~~G~~	1980-Present	0.625°×0.5°	2 months	(Gelaro et al. et al., 2017)
MERRA2-LND	MERRA2 Land	~~S~~, R, G	1980-Present	0.625°×0.5°	2 months	(Gelaro et al. et al., 2017)
MSWEP v.2.2	Multi-Source Weighted Ensemble Precipitation v.2.2	S, R, G	1979-2017	0.1°	Stopped	(Beck et al., 2
PERSIANN-CDR	Precipitation Estimates from Remotely Sensed Information using Artificial Neural Network and Climate Data Record	S, G	1983-2016	0.25°	6 month	(Ashouri et al
PERSIANN-CSS-CDR	PERSIANN-Cloud Classification System-CDR	S, G	1983-2016	0.04°	6 month	Sadeghi et al. Review)
TMPA-RT v.7	TRMM Multi-satellite Precipitation Analysis Real Time v.7	S	1998-Present	0.25°	1 day	(Huffman et a 2010)
TMPA-Adj v.7	TMPA Adjusted v.7	S, G	2000-Present	0.25°	3 months	(Huffman et a 2010)
WFDEI-CRU	WATCH Forcing Data methodology applied to ERA-Interim-Climatic Research Unit	R, G	1979-2016	0.5°	stopped	(Weedon et a
WFDEI-GPCC	WFDEI-Global Precipitation Climatology Center	R, G	1979-2016	0.5°	stopped	(Weedon et a

Figure 2 shows the mean annual precipitation derived from all considered P-datasets at their original spatial resolution over the 2001-2010 period. Although all P-datasets are able to represent the south-north precipitation pattern, they differ in terms of total amount and local patterns. These differences may reflect in the evaluation presented in this study. The blocky effect observed for GSMaP-Adj v.6 and PERSIANN-CSS-CDR may be attributed to the post adjustment of these P-datasets based on CPC and GPCP, respectively. Similar effects were observed for IMERG-F v.3, which was progressively removed along with the development of the updated version 4, 5 and now 6 (Satgé et al., 2018).

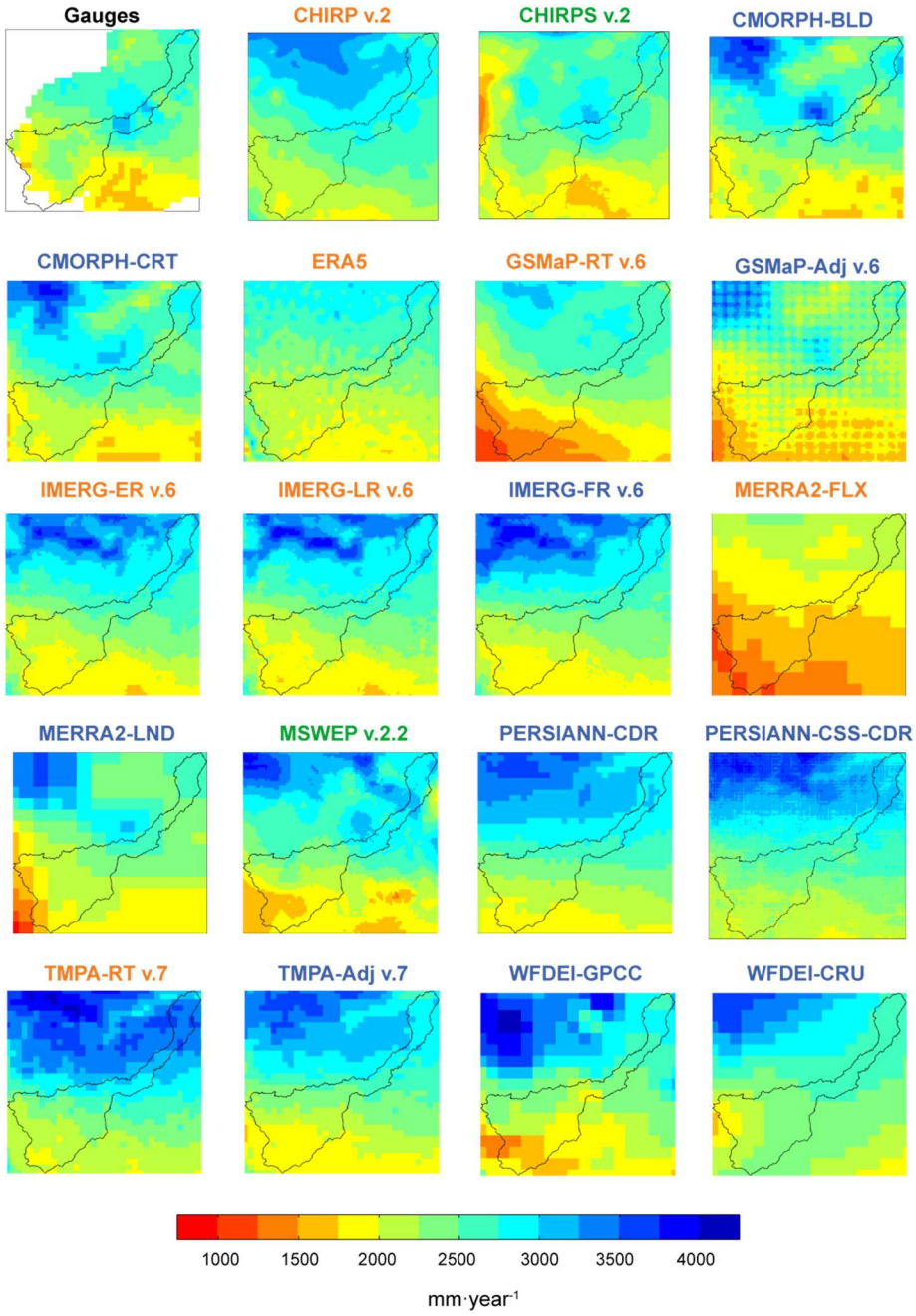


Figure 2. Mean annual precipitation maps from the considered P-datasets over the 2002-2010 period with different color name for P-datasets using gauge-based information (blue), P-datasets without gauge-based information (orange) and P-datasets using gauge, satellite and reanalysis-based information

2.4. GLEAM ETp

The potential evapotranspiration (ETp) from the Global Land Evaporation Amsterdam Model version 3a (GLEAM v.3a) (ETp) (Martens et al., 2017) is selected in this study as forcing data for the hydrological models.

GLEAM v.3a estimates are derived from the Priestley and Taylor equation using ERA-Interim as forcing data (net radiation and air temperature). Available at daily time step and 0.25° spatial resolution, this dataset was previously validated across the Altiplano (Satgé et al., 2019a) and used for hydrological modelling across the same region (Satgé et al., 2020) and West Africa (i.e. Dembélé et al., 2020).

2.5. Hydrological Models

Two lumped hydrological models, GR4j (Perrin et al., 2003) and HyMOD (Wagener et al., 2001), were selected for the analysis to avoid any influence from the models on the gridded P-dataset performance. Both models require ETp and P as forcing data.

GR4j was successfully used in studies conducted over the South American continent (Satgé et al., 2019 and 2020). However, in comparison to 35 hydrological models (including GR4j) HyMOD aims at performing relatively better in baseflow-dominated catchments without flashy streamflow behavior (Knoben et al., 2020) as observed over the Amazonian basin. Therefore, HyMOD should provide more accurate streamflow simulations and is considered for comparison.

HyMOD is a 5-parameter and 5-storage model (Figure 3a). It is based on a soil moisture storage in order to compute and separate the excess precipitation into quick and slow runoff. Quick runoff is routed through a Nash cascade of three identical linear reservoirs whereas slow runoff is routed in a parallel reservoir. The streamflow at the catchment outlet is computed by summing up quick and slow runoff.

GR4j is a 4-parameter and 2-storage model (Figure 3b). It first computes the amount of water available for runoff (i.e. effective precipitation) by using a production module. Then a routing function splits the effective precipitation into two components (90 % and 10%) through separate unit hydrographs (UH1 and UH2) in order to represent the delayed and direct runoff, respectively. Finally, the delayed and direct runoff are summed to get the catchment outlet streamflow.

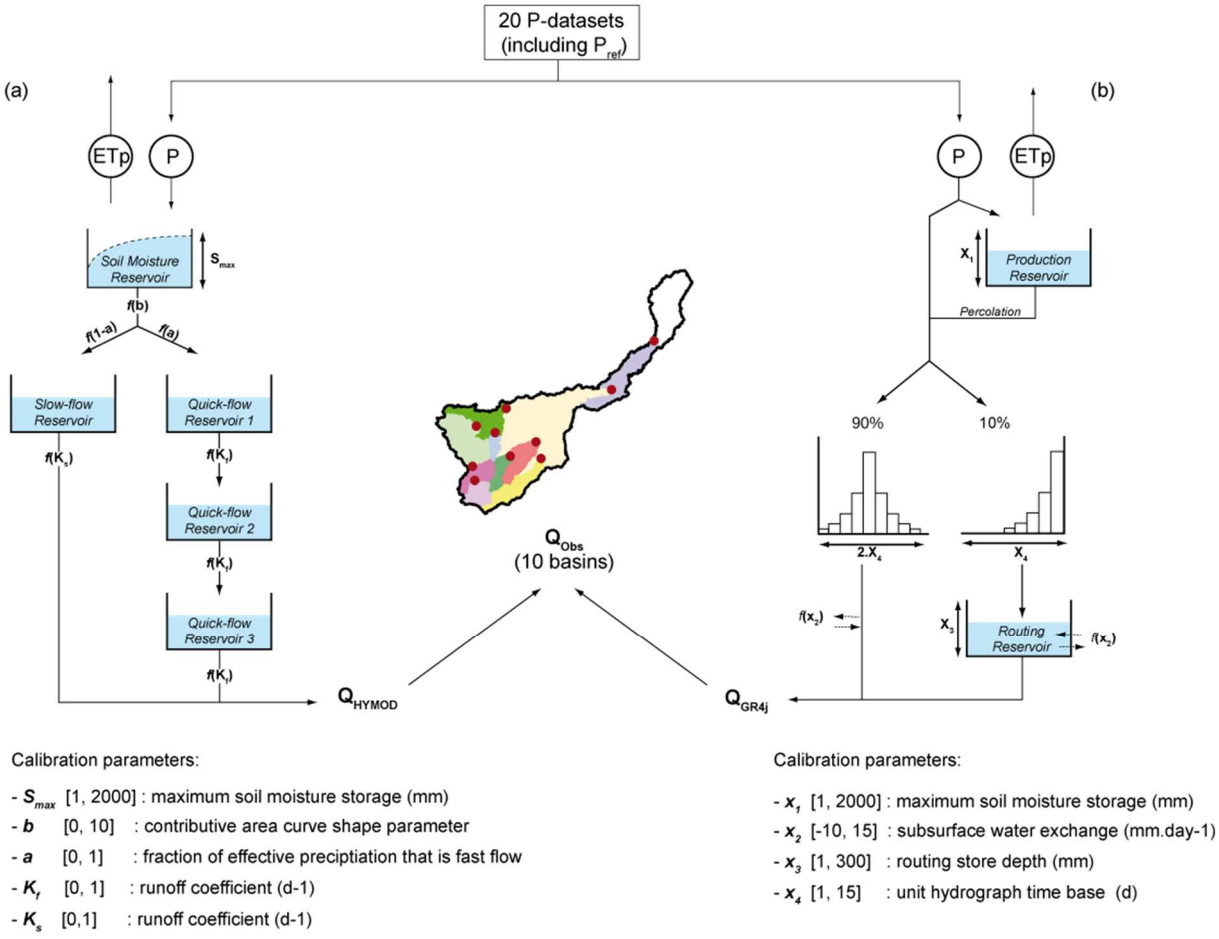


Figure 3. HyMOD (a) and GR4j (b) model description along with general workflow for P-dataset reliability in streamflow simulation.

3. Method

3.1. SPPs vs. Gauge Observations

To ensure a consistent comparison, all P-datasets (including the reference) were previously resampled from their original spatial resolution (table 1) to the 0.1° grid-cell size. This process relies on a bilinear mean (interpolation) for P-datasets with original spatial resolution <0.1° (>0.1°). Then, based on a weighted average for the grid-cells not fully included in the basins, the area-averaged daily P series were computed for each P-dataset and compared to the one derived from the reference (P_{ref}) using the Kling–Gupta efficiency (KGE) (Gupta et al., 2009) (eq. 1-4).

KGE is commonly used to assess P-dataset reliability as it combines the correlation coefficient (*R*), bias (*Bias*) and ratio of variability (V_r) between reference observations (in this case, the gauge observations) and evaluated P-datasets, all of which are relevant scores to look at for efficient management of water resources.

$$KGE = 1 - \sqrt{(R-1)^2 + (Bias-1)^2 + (Vr-1)^2} \quad (1)$$

206

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

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

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

207 Where μ and σ are the distribution mean and standard deviation, respectively; and s and o stand for the
208 estimate and the reference, respectively.

209 **3.2. SPPs vs Hydrological Modelling**

210 The open source Modular Assessment of Rainfall–Runoff Models Toolbox v.1.2 (MARRMOT) was used to run
211 the GR4j and HyMOD hydrological models (Knoben et al., 2019). Each of the two models was implemented in
212 the 10 basins under study. Note that each basin is treated independently of the other, including for nested
213 watersheds. This means that for each basin, the precipitation was aggregated over the whole area located
214 upstream of the streamflow gauges. MARRMOT uses the Nelder–Mead simplex algorithm for the objective
215 function optimization to calibrate the models. Generally, KGE and/or Nash Shuttle Efficiency score (NSE) are
216 used as objective functions for streamflow modelling (i.e. Fallah et al., 2020; Jiang and Bauer-Gottwein, 2019b;
217 Tarek et al., 2019; Zhang et al., 2020). A previous study assessing P-dataset reliability for streamflow modeling
218 has shown that P-dataset efficiency ranking remains equal independently of using KGE or NSE as objective
219 function (Satgé et al., 2020). Therefore, we only used KGE as objective function for this study.
220 For each precipitation input (P_{ref} and P-datasets), both models (i.e. GR4j and HyMOD) were calibrated over the
221 2001–2010 period (10 years) using the 2001 year as a spin-up period. As the objective of this study was to
222 assess P-dataset reliability for streamflow modelling rather than the hydrological model robustness, we did not
223 consider a validation period.

224 **4. Results**

225 **4.1. SPPs vs. Gauge Observations**

Figure 4 shows reliability of the P-datasets in representing daily and monthly average precipitations over the 10 considered basins.

At the daily time step, the 5 most efficient P-datasets are CMORPH-BLD v.1, GSMaP-Adj v.6, ERA5, CHIRPS v.2 and PERSIANN-CDR with KGE superior to 0.4 (0.56, 0.47, 0.43, 0.41 and 0.4, respectively). In terms of R, CMORPH-BLD v.1 and GSMaP-Adj v.6 present the highest scores of 0.71 and 0.75 while all other SPPs have R inferior to 0.6. In terms of Bias, most of the P-datasets present Bias close to one. However, MERRA2-FLX and GSMaP-RT v.6 (PERISANN-CDR, CHIRP v.2) strongly underestimate (overestimate) precipitation amounts. Finally, in terms of Vr, all P-datasets present higher precipitation variability than reference precipitation with Vr superior to 1. Among the top 5 P-datasets (in terms of KGE) ERA5 is the only one with Vr close to 1.

When considering the monthly time step, the P-dataset ranking changes to CHIRPS v.2, CMORPH-BLD v.1, ERA5, MERRA2-LND and PERSIANN-CDR with KGE value of 0.95, 0.87, 0.87, 0.87 and 0.86, respectively. CHIRPS v.2 ranked first (instead of CMORPH-BLD v.1 at daily time step), MERRA2-LND is now in the top 5 P-dataset and GSMaP-Adj v.6 is ranked 7^{th} with a KGE value of 0.76.

The different ranking observed at daily and monthly time step can partially be attributed to (i) the time step of the gauge-based datasets used for the precipitation adjustment and (ii) the 24-hours window used to calculate the daily precipitation amount (F. Satgé et al., 2020). For example, CMORPH-BLD v.1 is based on daily gauge-based datasets (CPC) while CHIRPS v.2 uses both daily and monthly gauge-based datasets (CHPclim + Gauges). As a result, CMORPH-BLD v.1 is the most efficient at the daily time step while it is CHIRPS v.2 at the monthly time step. Similarly, GSMaP-Adj v.6 uses daily gauge-based information and thus ranks in the top 5 daily P-datasets but no longer at the monthly time step. Additionally, the 24-hour time window used to retrieve daily precipitation does not systematically match that of the observed data. Therefore, the daily time step reliability might be influenced by this temporal mismatch (i.e. CHIRPS v.2, CMORPH-BLD v.1, PERSIANN-CDR-CSS). However, the monthly comparison is not affected, as the temporal mismatch is very low at this time step (few hours). It is important to note the difference in P-dataset performance at daily and monthly time step since the choice of the P-dataset may differ according to the intended use.

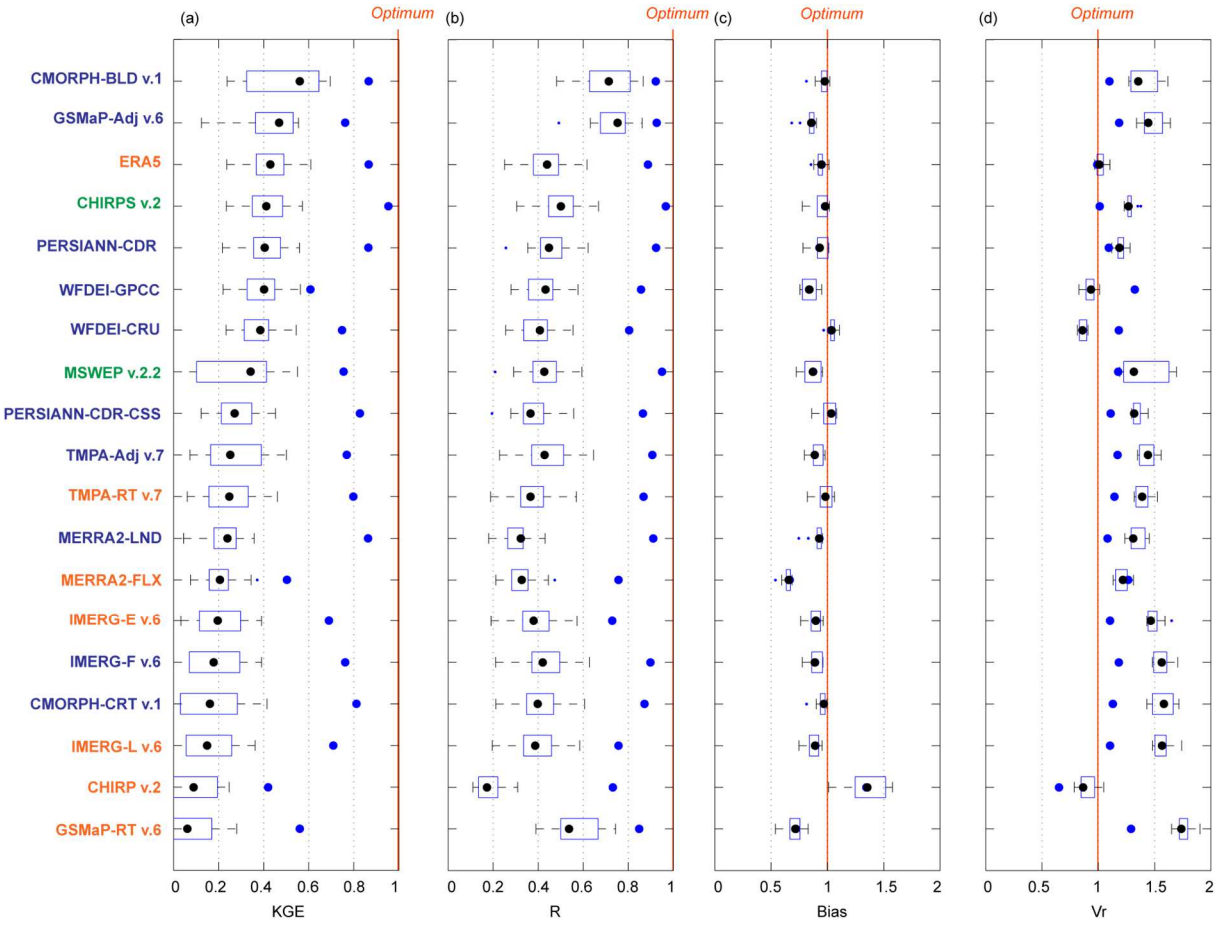


Figure 4. P-dataset reliability to represent the regional average precipitation of the 10 considered basins in terms of KGE (a), R (b), Bias (c) and Vr (d). The right and left edges of the boxes represent the 25th and 75th percentile values, respectively. The P-datasets are sorted from the most (top) to the least (bottom) efficient in terms of KGE and for the daily time step with different color name for P-dataset using gauge-based information (blue), P-dataset without gauge-based information (orange) and P-datasets using gauge, satellite and reanalysis-based information. Black and Blue dots represent the median value obtained at daily and monthly time step, respectively.

Figure 5 shows the KGE along with the average absolute difference between P-dataset and P_{ref} (in terms of absolute value simplicity purpose) value obtained from all P-datasets and for all considered catchments. Only KGE and Bias are represented because Bias was found to have the dominant influence on KGE values (Satgé et al., 2020).

In general, the reliability of most of the considered P-datasets presents a south-north gradient with an increasing KGE value from south to the north. Along this gradient, the surface extent of the catchments increases showing an improvement between P–datasets and P_{ref} observations (grid-cell spatial average vs. point observation).

While CHIRPS v.2 and MSWEP v.2.2 include satellite, reanalysis and gauge-based estimates, they do not provide the most reliable daily precipitation estimates across the considered basins (Fig. 5). However, when considering the monthly time step, CHIRPS v.2 is the most reliable P-dataset in terms of KGE, and MSWEP v.2.2 is the most reliable in terms of R (Fig. 4 a and b). This discrepancy between daily and monthly reliability for CHIRPS v.2 and MSWEP v.2.2 may partially be related to the time step of the gauge-based information they used. Indeed CHIRPS v.2 is based on monthly gauge-based estimates (CHPclim) and therefore performs better at monthly than daily time step. This is in line with previous results obtained across West Africa showing the influence of the gauge-based information time step over the P-dataset reliability (Satgé et al., 2020). Similarly, MSWEP v.2.2 uses both daily and monthly gauge-based estimates instead of only daily gauge-based estimates, which may attenuate the daily dataset influence in MSWEP precipitation estimates. Additionally, the daily-gauge based datasets used in MSWEP v.2.2 (Rain gauges from different sources) differ from the one used by the other P-datasets (CPC). The gauge-based information used by MSWEP v.2.2 might be less reliable than CPC (in the Juruá basin) and accounts for the relatively lower reliability of MSWEP at the daily time step.

Finally, the improvement observed from CHIRP v.2 (GSMaP-RT v.6) to CHIRPS v.2 (GSMaP-Adj v.6) is notable with an increase (decrease) of the KGE (Bias) for all considered basin (Figure 5). However, taking into account gauge-based precipitation estimates does not systematically lead to such an improvement. Indeed, considering the IMERG-E, -L and –F v.6 datasets, the daily KGE and Bias patterns are very similar regardless of the IMERG versions (Figure 5). Similarly, MERRA2-FLX and –LND present a similar KGE pattern and the improvement from MERRA2-FLX to MERRA2-LND is limited to a noticeable decrease in Bias. Despite a slight increase in terms of KGE from TMPA-RT to TMPA-Adj v.6, the inclusion of gauge-precipitation estimates also led to a slight increase of Bias value.

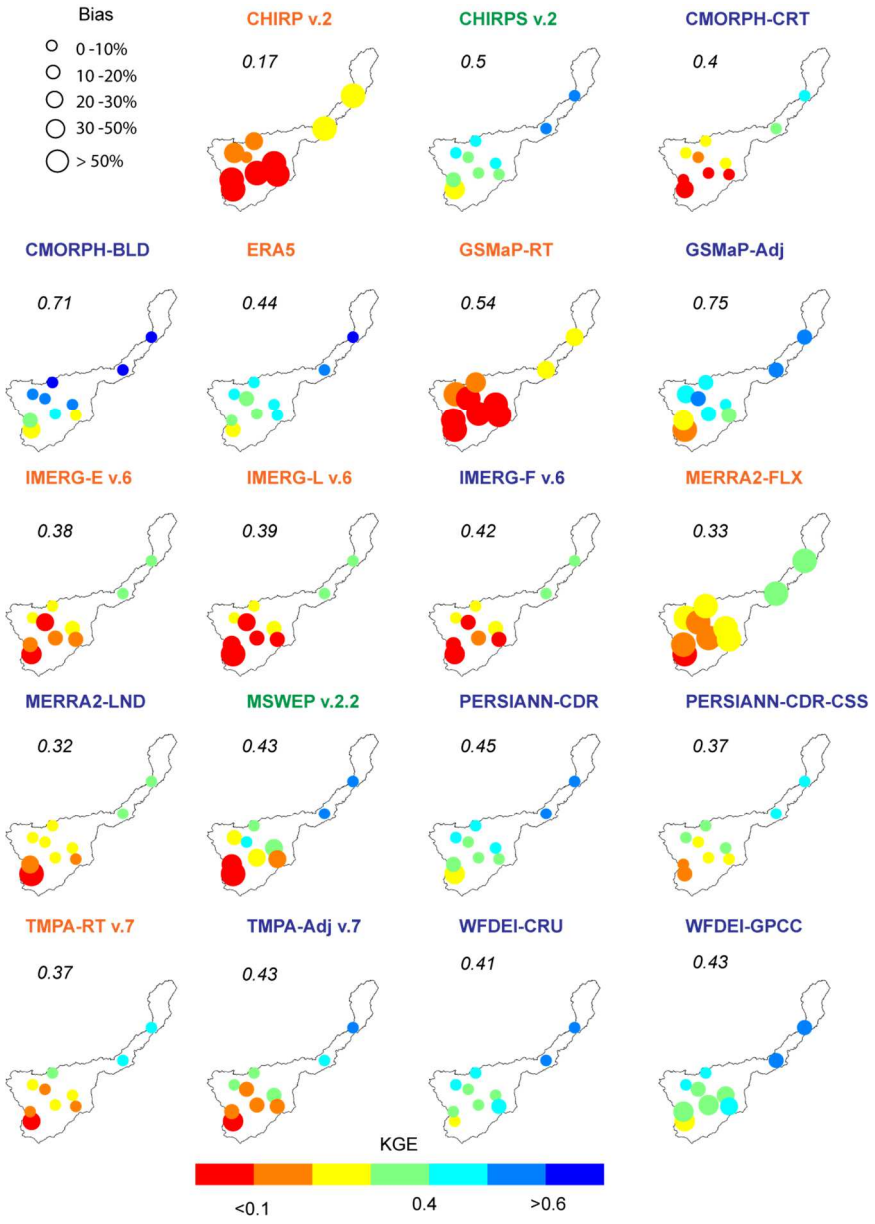


Figure 5 Daily P-dataset KGE and Bias (absolute difference between P-dataset and P_{ref} for simplicity purpose) for the 10 basins with different color name for P-dataset using gauge-based information (blue), P-dataset without gauge-based information (orange) and P-datasets using gauge, satellite and reanalysis-based information. Italic values indicate the median R value obtained from the 10 basins.

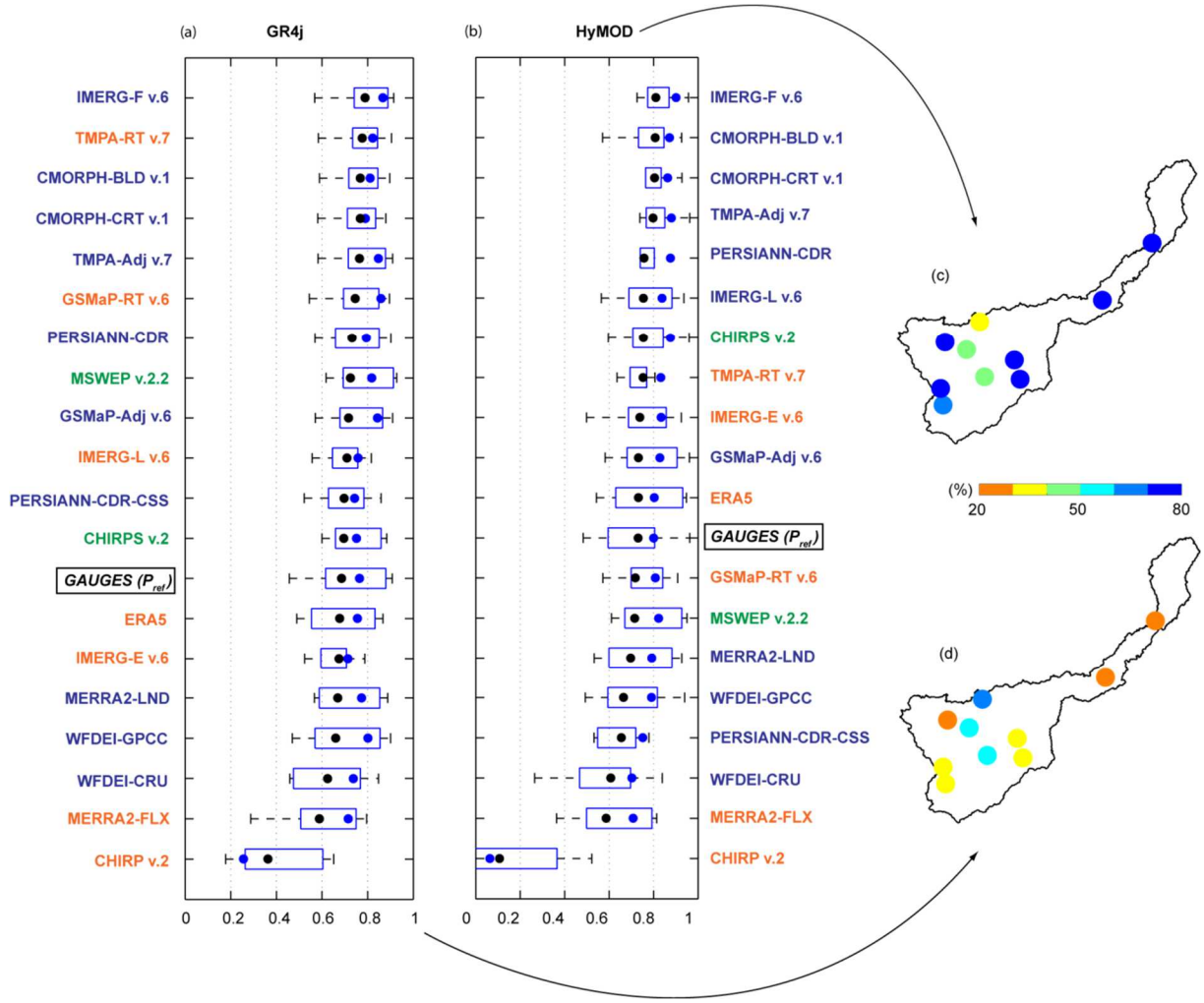
4.2. SPPs vs Hydrological Modelling

Figure 6 depicts the performance of the 19 P-datasets and P_{ref} for streamflow modelling using GR4j and HyMOD. For both models, IMERG-F v.6 as forcing data provides the most efficient streamflow simulation with median KGE of 0.79 and 0.81. Conversely, CHIRP v.2 provides the less efficient streamflow simulation with KGE value of 0.36 and 0.11 using GR4j and HyMOD, respectively. Interestingly, P_{ref} is not the most efficient option for streamflow modelling over the Juruá basin. Actually, half of the considered P-datasets provide more

realistic streamflow simulations. This could be related to the few gauges available across the region, which are not able to consistently represent the local precipitation pattern. Different observations could be achieved across basins with higher gauge density.

When comparing daily and monthly streamflow simulations, for both models, the streamflow simulations are closer to the observations at monthly time step than daily. Interestingly, the P-dataset ranking efficiency differs between daily and monthly simulations. Indeed, IMERG-F v.6, TMPA-RT v.7 and CMORPH-BLD v.1 (IMERG-F v.6, CMORPH-BLD v.1 and CMORPH-CRT v.1) are the top 3 P-datasets for GR4j (HyMOD) at daily time step whereas IMERG-F v.6, GSMaP-RT v.6 and TMPA-Adj v.7 (IMERG-F v.6, TMPA-Adj v.7 and CHIRPS v.2) are the top 3 P-datasets at the monthly time step. The better performance of TMPA-Adj v.7, PERSIANN-CDR and CHIRPS v.2 at the monthly time step can be partially explained by the adjustment of the precipitation estimates using monthly precipitation datasets. Despite a precipitation adjustment based on monthly data (GPCC), IMERG-F v.6 is the most efficient P-dataset at both daily and monthly time step.

Finally, when comparing GR4j and HyMOD performances, the use of the HyMOD model led to a slightly more accurate streamflow simulation than GR4j (Figure 6 c and d) at all basin outlets, except at the 3 outlets located in the west-central region. Based on this observation the following analysis only considers HyMOD model.



316

317 *Figure 6. P-dataset reliability in terms of KGE to represent the streamflow at the outlet of the 10 considered basins using*
318 *GR4j (a) and HyMOD (b). The right and left edges of the boxes represent the 25th and 75th percentile values, respectively.*
319 *For each model, the P-datasets are sorted from the most (top) to the least (bottom) efficient in terms of KGE. HyMOD (c) and*
320 *GR4j (d) frequency to reach the highest KGE considering all P-datasets at daily time step and for all basins. The blue dots*
321 *represent the median KGE obtained at the monthly time step.*

322 Figure 7 shows the P-dataset reliability for daily streamflow modeling at the outlet of the 10
323 considered basins. All P-datasets (including P_{ref}) are more reliable for the two basins located eastward (Figure
324 7), as the hydrograph is smoother (Figure 1) and therefore easier to simulate. Actually, the P-datasets are less
325 reliable for the streamflow simulation of the smallest basins located westward, as their hydrographs are much
326 more reactive to daily precipitation variability (Figure 1). This trend is clearly observable in Figure 8, which
327 shows that the fit between modelled and observed streamflow improves with basin area.

Interestingly, the incorporation of a precipitation post-adjustment in CHIRPS v.2 and IMERG-F v.6 considerably increases their reliability for daily streamflow modeling compared to their non-adjusted version CHIRP v.2 and IMERG-E and –L v.6, respectively. Indeed, CHIRP v.2 considerably overestimates streamflow for both high and low flow periods. This is even more significant for the small catchments under study (Figure 8). Regarding the IMERG datasets, the overestimation of IMERG-E and-L v.6 streamflow's observed during the high flow period is consistently corrected in IMERG-F v.6, showing the benefit of including gauge-based precipitation estimates. To a lesser extent, a similar improvement is observed for TMPA-Adj v.7 and MERRA2-LND compared to their no-adjusted versions TMPA-RT v.7 and MERRA2-FLX, respectively. Indeed, the TMPA-RT v.7 (MERRA2-FLX) streamflow overestimation during the high flow (low flow) period (Fig. 8) is well corrected in TMPA-Adj v.7 (MERRA2-LND) (Fig .8). Interestingly, KGE values are similar for GSMaP-RT and –Adj v.6 (Fig. 7). However, when examining the streamflow series, GSMaP-RT v.6 presents a "time lag" consisting in a streamflow overestimation (underestimation) for the first (second) part of the hydrological cycle (Fig. 8). Thus, the positive and negative bias (i.e. streamflow over and underestimation) are balanced, leading to similar KGE value to that of GSMaP-Adj v.6. However, despite similar KGE values, GSMaP-Adj v.6 streamflow simulation are much more reliable as it removes the "time lag" observed for GSMaP-RT v.6 (Fig. 8).

Interestingly, even if CHIRPS v.2 and MSWEP v.2.2 include satellite, reanalysis and gauge-based precipitation estimates, they do not provide the most reliable streamflow simulation across the considered basins (Figure 6). As reported in the previous section, this might be partially related to the relatively lower reliability of the gauge-based precipitation used in CHIRPS v.2 and MSWEP v.2.2 when compared to the one used by the other P-datasets.

Finally, even if the P-datasets including gauge-based precipitation estimates (IMERG-F v.6, CMORPH-BLD, CMORPH-CRT and TMPA-Adj v.7) generally outperform the other P-datasets (Figure 6) they are not systematically the most reliable for all considered basins (Figure 7). This shows that P-datasets are subject to spatial "inconsistency" even in a relatively restricted region such as the Juruá basin. This is consistent with previous results having highlighted the importance of evaluating P-datasets, in order to ensure that the most reliable P-dataset is selected (Satgé et al. 2020, Pius et al., 2021 Beck et al., 2019). Finally, for long term analysis, with approximately 40 years of observations, PERSIANN-CDR appears as the most interesting option to follow daily streamflow dynamics across the Juruá basin. Despite a higher spatial resolution (0.04°), the recently released PERSIANN-CSS-CDR is not as reliable as the lower spatial resolution (0.25°) PERSIANN-CDR.

Indeed, even for the smallest basins considered in our study, the PERSIANN-CSS-CDR product shows worse performance than PERSIANN-CDR. Moreover, CHIRPS v.2, which presents the second best spatial resolution among P-datasets, does not provide the better score across small basins (Figure 7).

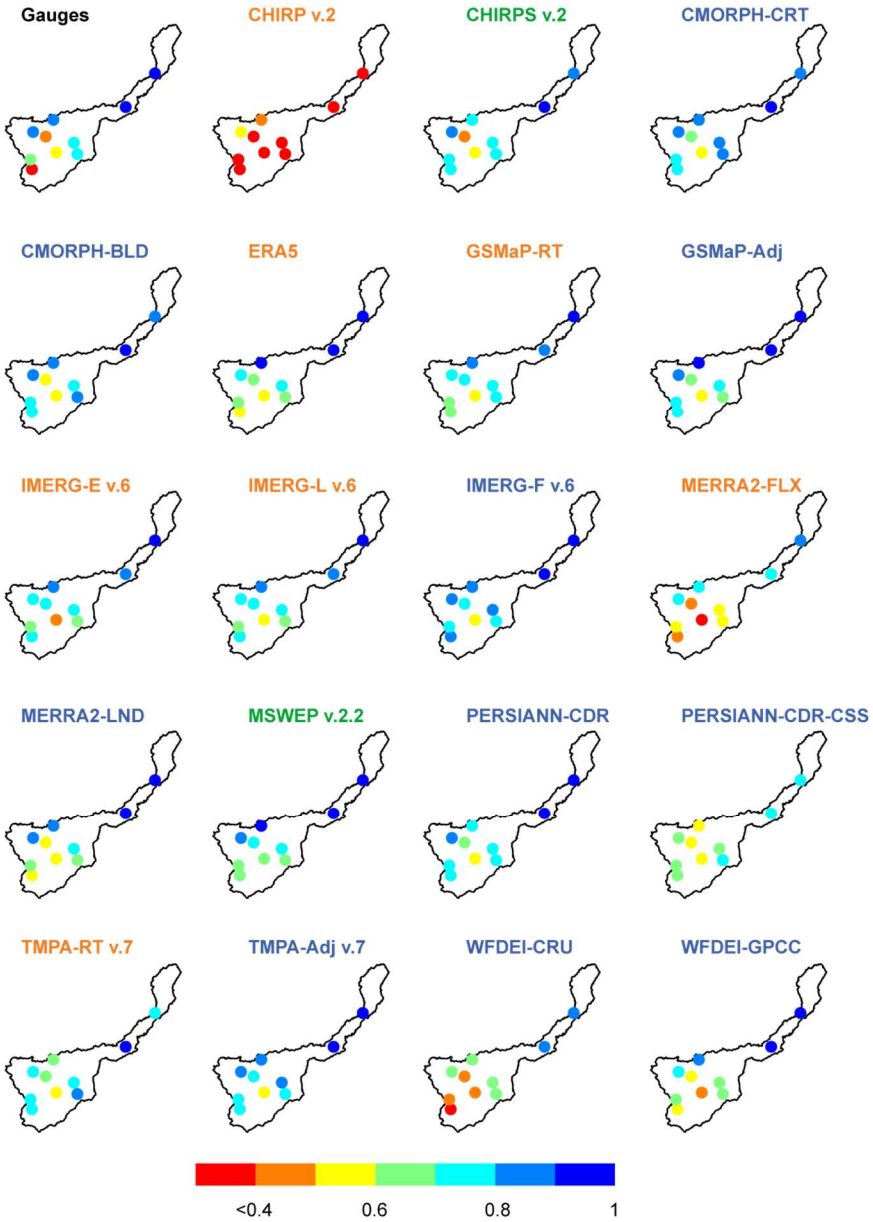


Figure 7. P-dataset reliability for HyMOD's daily streamflow simulation expressed in the form of KGE at the outlet of the 10 considered basins with different color name for P-dataset using gauge-based information (blue), P-dataset without gauge-based information (orange) and P-datasets using gauge, satellite and reanalysis-based information. Italic values indicate the median correlation value obtained from the 10 basins

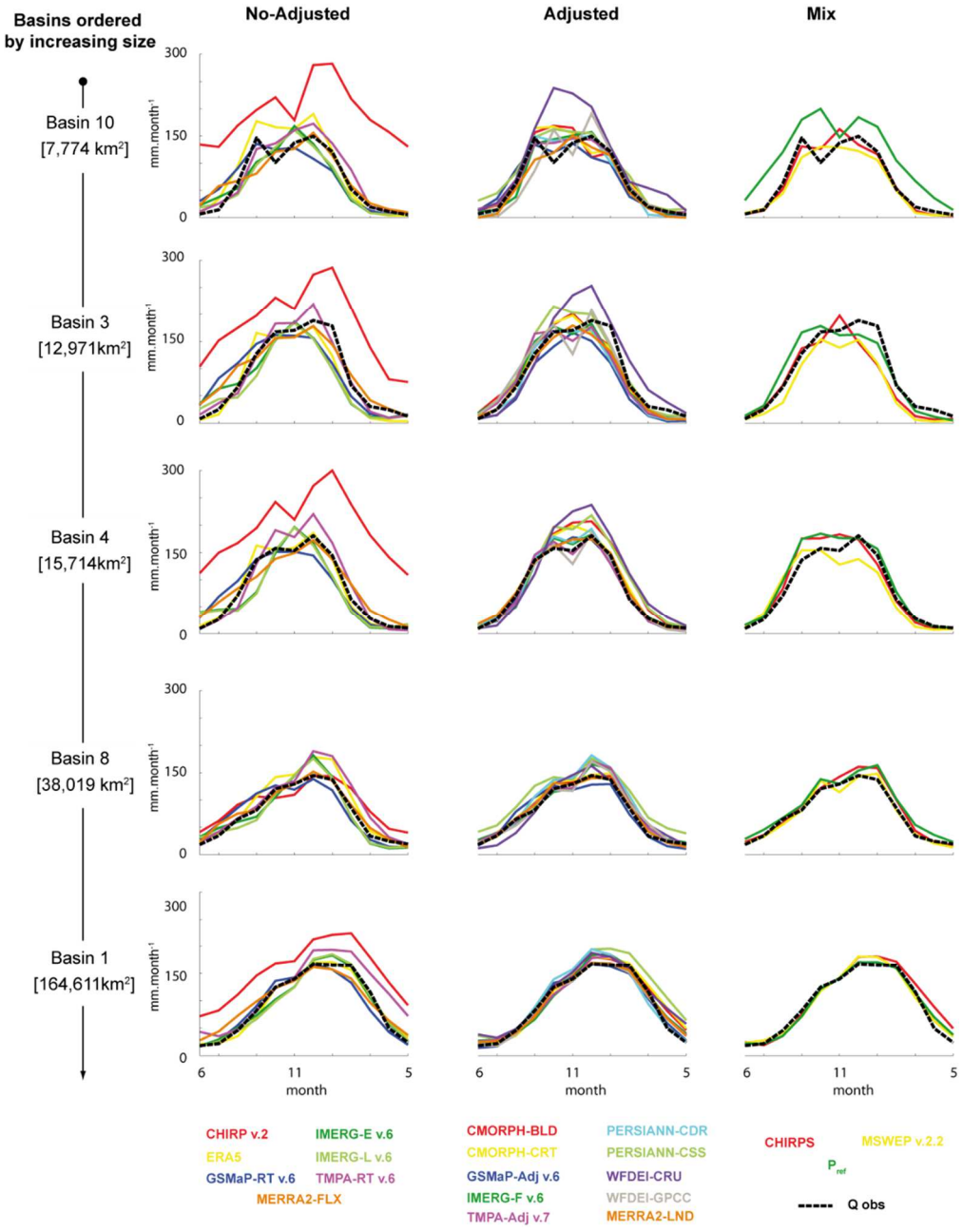


Figure 8. HyMOD's mean monthly streamflow for the 2002-2010 period obtained from all considered P-dataset and for 5 out of the 10 basins. The basins are sorted from the largest (bottom) to the smallest (top) area.

5. Discussion

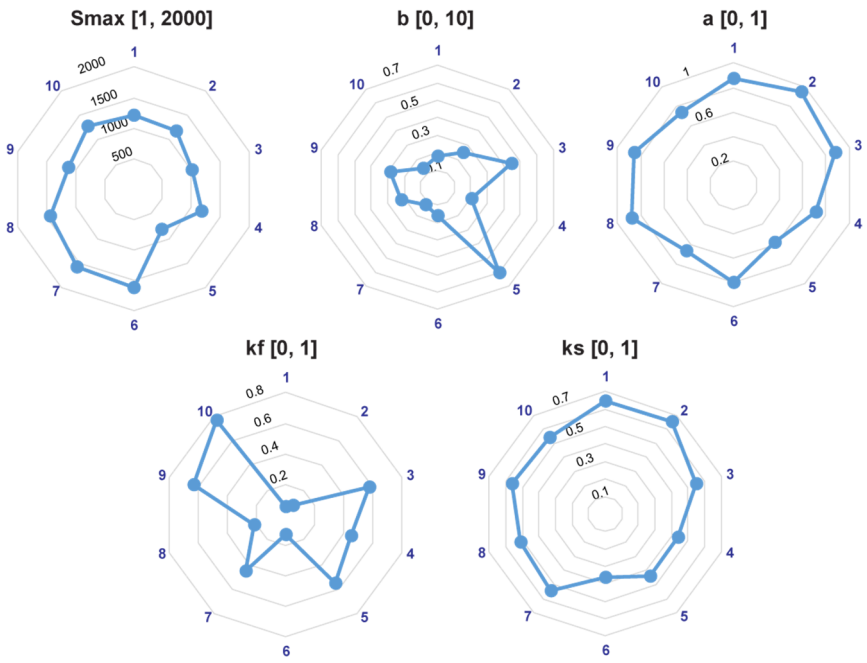
5.1 Hydrologic model parameter consistency

Figure 9 shows the median parameters values derived from all P-datasets and for all considered basins. The values of the parameters "b" and "Smax" ($0.1 < b < 0.6$; $808 < Smax < 1606$) which control the soil moisture storage capacity are relatively similar among the watersheds, except for catchment 5 which has a lower storage capacity. Parameter "a" divides the quick and slow flows, its value varies between 0.58 and 0.94. Parameter

374 "Kf" controls the exchange between the 3 quick flow storages and its value varies between, 0.05 up to close
375 0.77. The highest "Kf" values are found for the headwater basins (smallest one). "Ks" that controls the slow
376 flow ranges from 0.36 to 0.65.

377 Overall, all parameter values remain within the range advised by model developers and relatively
378 constant among basins. The most variable parameter is "Kf" reflecting the slope change across the study area.
379 Moreover, the parameter "b" remains low (<0.7) while the parameter "a" is in its upper half range (>0.5) for all
380 considered basins especially the headwater ones (basins 5, 6, 7, 8). This means that precipitation is more likely
381 to contribute to quick flow than slow flow. This feature drastically differs with GR4j model and might explain
382 that HyMOD performs better than GR4j across the Juruá basin. Indeed, in GR4j 90% and 10% of the
383 precipitation are arbitrarily routed as slow and quick flow, respectively (fig. 3) whereas routing is calibrated in
384 HyMOD. This parametrization favors quick flow in the Juruá basin.

385



386

387 *Figure 9. HyMOD's mean parameter values obtained from all considered P-datasets and for all considered basins*

388 **5.2. P-dataset reliability for different basins' sizes**

389 A previous study led across the Mediterranean region has shown that P-datasets are less reliable for small
390 basin areas (Camici et al., 2018). As a similar pattern emerged from Figures 7 and 8, Figure 10 shows the P-
391 dataset reliability for the different basins and according to their surface areas. P-dataset reliability grows with

the size of the considered basins. Actually, KGE superior to 0.8 is observed for basin areas superior to approximately 16,000 km^2 (Figure 10). Precipitation aggregation over a large area minimizes the influence of local precipitation variation. With a spatial resolution of several km^2, the P-datasets are not suitable for representing local precipitation dynamics. This drawback is less significant when considering larger basin and could explain why P-dataset reliability increases with the basin area. Obviously, hydrological models are also more efficient in representing smooth hydrographs as observed over large basins than dynamic ones as observed over smaller basins. It is worth mentioning that KGE values do not reflect the absolute P-dataset reliability because lumped models such as HyMOD are not the best approach for large basins. These values are only given to compare P-dataset reliability with each other.

Note that in most cases, the best streamflow simulation is always achieved using a P-dataset instead of P_{ref} (Figure 10). However, this statement is only true for the region under study because only few gauges are available to develop the P_{ref}. An opposite conclusion could be drawn if more rain gauges were available across the region. According to the actual context and for the time being, P-datasets offer a great opportunity for streamflow simulation.

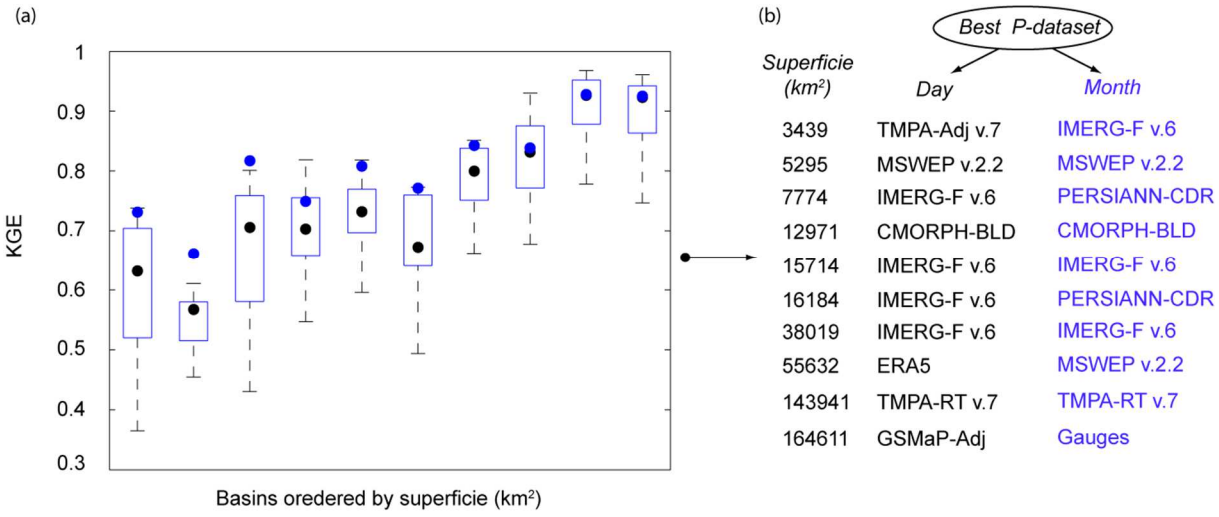


Figure 10. Overall P-dataset reliability for streamflow simulation (HyMOD) for different basins areas sorted according to increasing size (a), with the best P-datasets at daily and monthly time step (b). Black (blue) dots represent the median KGE obtained at the daily (monthly) time step.

5.1. Gauges vs. hydrological modeling assessment

411 Daily KGE is low for most of the considered P-datasets (KGE <0.4 for 14 out of 19 P-dataset) when comparing to
412 P_{ref}. Indeed, P-datasets provide spatially average precipitation observation across grid-cell superficial extent
413 ranging from 16 km^2 (PERSIANN-CDR-CSS) to 3,125 km^2 (MERRA2-FLX, -LND) while rain gauges provide point
414 measurement at the exact gauge location. Across tropical regions, precipitation presents high variation in space
415 so that many precipitation events captured at the grid-cell level might be unregistered at the gauge level.
416 Therefore, the relationship (i.e. KGE in this study) between the observations made by the P-datasets and the
417 rain gauges improves with the number of rain gauges used to represent the rainfall of the considered grid-cell
418 (Salles et al., 2019, Tang et al., 2018). Similar discrepancy between the basin spatially average measurement
419 observed by the P-datasets and P_{ref} occurs in this study and explains the low relationship between P-datasets
420 and P_{ref}. This discrepancy is less pronounced at the monthly time step as monthly precipitation generally shows
421 lower spatial variation than daily one.

422 The mismatch between P-dataset grid-cell measurement and point gauge observation clearly shows the limit of
423 such P-dataset reliability evaluation (especially in the case of scarce gauge network such as the Juruá amazon
424 basin) and the necessity of alternative methods such as the one using hydrological modelling.

425 Indeed, despite being a point measurement, the streamflow observation is directly connected to the spatially
426 average precipitation of its corresponding basin. With spatially average observations (grid-cell), P-datasets are
427 likely to provide more reliable basin spatially average precipitation estimates than P_{ref}, which relies on the
428 interpolation of scarce gauge network.

429 Actually, the streamflow simulation obtained using P_{ref} are less reliable than the one obtained using most of the
430 P-datasets (10 and 12 out of 19 if using HyMOD or GR4j model, respectively). However, P-dataset reliability
431 based on hydrological modelling presents some limits such as : (1) the influence of the model parameters
432 adjustment which can compensates part of the P-dataset uncertainty, (2) the model itself as different
433 conclusions should be drawn according to the considered models (Figure 6), and (3) the impossibility of
434 assessing P-dataset at very local scale (grid-cell to grid-cell) as precipitation are averaged at the basin scale.

435 Despite the advantage of using streamflow modelling to assess P-dataset reliability, in many region P-dataset
436 reliability can only be assessed through comparison with rain gauge estimates as streamflow data are not
437 available or because the basin hydrological dynamic is highly influenced by human activities (withdraw,

agriculture, cities, dams etc..). In this context, it is crucial to use statistical indices leading to the most relevant conclusion that would have been drawn using a streamflow modelling approach.

A recent study led across Europe has shown that common statistic indicators used to assess P-dataset reliability against rain gauge observations do not reflect the true P-dataset potential for streamflow modeling (Camici et al., 2020). The authors show that Relative Root Mean Square error (RRMSE eq.4) and Absolute relative Mean Error (AME eq. 5) better reflect P-dataset potential than KGE. In this context, we follow the same approach as Camici et al. (2020) to verify this statement across the region of interest.

$$RRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(s_i - o_i)^2}}{\frac{1}{n}\sum_{i=1}^{n} o_i} \quad (5)$$

$$AME = \frac{\frac{1}{n}\sum_{i=1}^{n}(s_i - o_i)}{\frac{1}{n}\sum_{i=1}^{n}(o_i)} \quad (6)$$

Where s and o indicate the estimate and the reference, respectively

Figure 11 compares the KGE obtained for streamflow modeling (KGE_Q) with KGE, RRMSE and AME obtained when comparing basin average precipitation derived from the P-dataset with the one derived from P_{ref}. Additionally, the median KGE_Q obtained for successive classes (with 0.1 step) of KGE, RRMSE and AME is also represented. Finally, we also consider the sum of RRMSE and AME in the analysis (Figure 11d).

KGE_Q is less correlated to KGE (CC=0.4) than to RRMSE (CC=-0.46) and AME (CC=-0.53). This is in line with the results obtained across Europe (Camicci et al. 2020). The results suggest that in the absence of streamflow observation, P-dataset assessment for hydrological modelling should use RRMSE and AME metrics instead of KGE. Actually, it is even more adapted to consider both RRMSE and AME as the correlation with KGE_Q is higher (CC= -0.56).

Across the Juruá basin, a threshold of 1.1 (RRMSE) and/or 0.1 (AME) and/or 1.1 (RRMSE+AME) leads to very good P-dataset performance for streamflow modelling (KGE_Q >= 0.75). It is worth to mention here that these thresholds could vary according to the region of interest and basin size as P-dataset reliability varies in

space (i.e. Satgé et al., 2020, Beck et al., 2019) and depends on the basin area (i.e. Camici et al., 2018 and Figure 10).

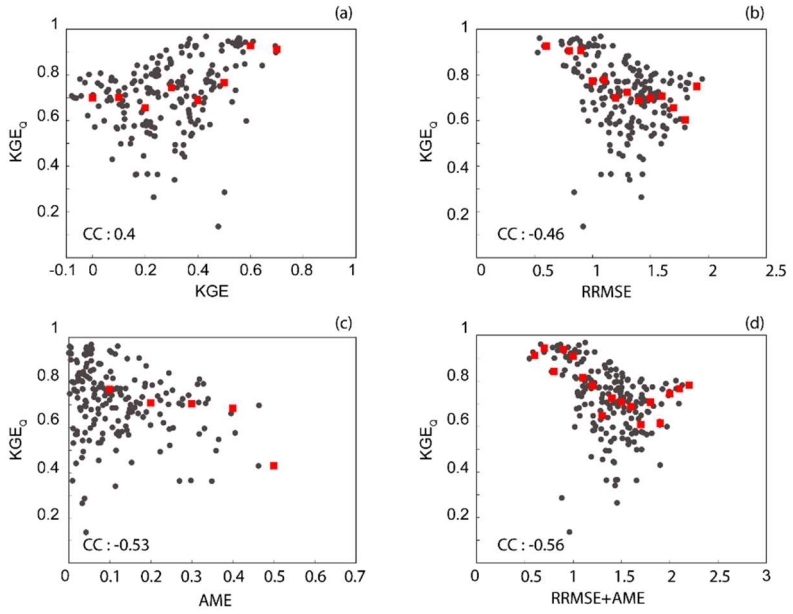


Figure 11. P-dataset statistic score when compared to gauges or streamflow with the red square representing the median KGE_Q obtained for successive classes (with 0.1 step) of KGE (a) RRMSE (b), AME (c) and RRMSE+AME (d)

Conclusion

This study assesses for the first time the reliability of 19 P-datasets for streamflow modelling across the Juruá basin, located in the Amazon region, using two lumped hydrological models over 10 basins. The results can be summarized as follows:

- When compared to rain gauge observations, all P-datasets are more reliable at the monthly than daily time step. However, the reliability of the P-datasets at the daily time step does not systematically reflect its efficiency at the monthly time step. In fact, the observed ranking at daily and monthly time step differs. This is partially explained by the gauge-based estimates time step used for the post adjustment. Using daily (monthly) gauge-based estimates for adjustment improves daily (monthly) P-dataset reliability.
- P-dataset reliability increases with the basin surface area. This is partially explained by (i) precipitation aggregation at the basin scale which minimizes P-dataset limits to represent local precipitation variations and (ii) the hydrograph smoothness for large basins which is easier to simulate by the hydrological models.
- In the absence of streamflow data to assess P-dataset reliability for streamflow modelling, the comparison of P-dataset amounts to gauge observations should be based on RRMSE and AME statistic metric rather than KGE. Indeed, RRMSE and AME better reflect P-dataset reliability for streamflow modelling than KGE.
- Among the P-datasets under study, IMERG-F v.6 is the most reliable one for daily and monthly streamflow modelling. However, IMERG-F v.6 only covers the 2000-present period. Therefore, users interested in longer temporal coverage should consider using PERSIANN-CDR, CHIRPS v.2, ERA5 and MSWEP v.2.2 (in this order) which cover the last 40 years.
- With shorter delay between observation measurement and availability of the data, only the satellite-based P-datasets TMPA-RT v.7, IMERG-L v.6 and GSMaP-RT v.6 provide satisfactory streamflow simulation. These P-datasets should be considered in the case of “near real time” observation requirement.

- Despite a finer spatial resolution than PERSIANN-CDR, PERSIANN-CSS-CDR does not provide more reliable precipitation estimates across the Juruá basin.

Overall, the results obtained in this study provide very useful information about the application of an exhaustive sample of available P-datasets (19) in simulating river discharge at basin scale across the Amazon region. This should be very helpful for data users facing the selection of the best P-datasets for long term and near real time applications.

Aknowledgments

This research was achieved in the framework of the BONDS project, funded through the 2017-2018 Belmont Forum and BiodivERsA joint call for research proposals, under the BiodivScen ERA-Net COFUND programme, and with the funding organisations French National Research Agency (ANR), São Paulo Research Foundation (FAPESP), National Science Foundation (NSF), the Research Council of Norway and the German Federal Ministry of Education and Research (BMBF).

The authors are grateful to P-dataset and GLEAM dataset providers, to Xavier et al. 2015 and 2017 for in situ precipitation and to CAMELS-BR (Chagas et al., 2020) for streamflow observations.

507 **References**

508 Adler, R.F., Gu, G., Huffman, G.J., 2012. Estimating Climatological Bias Errors for the Global Precipitation
509 Climatology Project (GPCP). J. Appl. Meteorol. Climatol. 84–99. https://doi.org/10.1175/JAMC-D-11-
510 052.1

511 Adler, R.F., Huffman, G.J., Chang, A., Ferraro, F., Xie, P.-P., Janowiak, J., Rudolff, B., Schneider, U., Curtis, S.,
512 Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E., 2003. The Version-2 Global Precipitation
513 Climatology Project (GPCP) Monthly Precipitation Analysis (1979 – Present). J. Hydrometeorol. 1147–
514 1167.

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

518 Beck, H.E., Pan, M., Roy, T., Weedon, G.P., Pappenberger, F., Dijk, A.I.J.M. Van, Huffman, G.J., Adler, R.F.,
519 Wood, E.F., 2019. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the
520 CONUS. Hydrol. Earth Syst. Sci. 207–224. https://doi.org/https://doi.org/10.5194/hess-23-207-2019

521 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.
522 MSWEP V2 global 3-hourly 0.1° precipitation methodology and quantitative assessment. Bull. Am.
523 Meteorol. Soc. 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

524 Becker, A., Finger, P., Rudolf, B., Schamm, K., Schneider, U., Ziese, M., Precipitation, G., Centre, C.,
525 Wetterdienst, D., 2013. A description of the global land-surface precipitation data products of the Global
526 Precipitation Climatology Centre with sample applications including centennial (trend) analysis from
527 1901–present 71–99. https://doi.org/10.5194/essd-5-71-2013

528 Brocca, L., Massari, C., Pellarin, T., Filippucci, P., Ciabatta, L., Camici, S., Kerr, Y.H., Fernández-Prieto, D., 2020.
529 River flow prediction in data scarce regions: soil moisture integrated satellite rainfall products
530 outperform rain gauge observations in West Africa. Sci. Rep. 10, 1–14. https://doi.org/10.1038/s41598-
531 020-69343-x

532 Camici, S., Ciabatta, L., Massari, C., Brocca, L., 2018. How reliable are satellite precipitation estimates for driving

hydrological models: A verification study over the Mediterranean area. J. Hydrol. 563, 950–961. https://doi.org/10.1016/j.jhydrol.2018.06.067

Camici, S., Massari, C., Ciabatta, L., Marchesini, I., Brocca, L., 2020. Which rainfall score is more informative about the performance in river discharge simulation? A comprehensive assessment on 1318 basins over Europe. Hydrol. Earth Syst. Sci. 24, 4869–4885. https://doi.org/10.5194/hess-24-4869-2020

Chagas, V.B.P., L. B. Chaffe, P., Addor, N., M. Fan, F., S. Fleischmann, A., C. D. Paiva, R., Siqueira, V.A., 2020. CAMELS-BR: Hydrometeorological time series and landscape attributes for 897 catchments in Brazil. Earth Syst. Sci. Data 12, 2075–2096. https://doi.org/10.5194/essd-12-2075-2020

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

Dembélé, M., Hrachowitz, M., Savenije, H.H.G., Mariéthoz, G., Schaefli, B., 2020. Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets. Water Resour. Res. 56, 1–26. https://doi.org/10.1029/2019WR026085

Dembele, M., Schaefli, B., Van De Giesen, N., 2020. Suitability of 17 gridded rainfall and temperature datasets for large-scale hydrological modelling in West Africa. Hydrol. Earth Syst. Sci. 24, 5379–5406. https://doi.org/10.5194/hess-24-5379-2020

Fallah, A., O, S., Orth, R., 2020. Climate-dependent propagation of precipitation uncertainty into the water cycle. Hydrol. Earth Syst. Sci. Discuss. 1–19. https://doi.org/10.5194/hess-2019-660

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, 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, 4302–4315. https://doi.org/10.1002/joc.5086

Funk, Chris, 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

559 environmental record for monitoring extremes 1–21. https://doi.org/10.1038/sdata.2015.66

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

562 Gebregiorgis, A.S., Hossain, F., 2013. Understanding the Dependence of Satellite Rainfall Uncertainty on
563 Topography and Climate for Hydrologic Model Simulation. IEEE Trans. Geosci. Remote Sens. 51, 704–718.
564 https://doi.org/10.1109/TGRS.2012.2196282

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

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

573 Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic
574 observations - the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711

575 Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorroshian, S., Tan, J., Xie, P.,
576 2018. Algorithm Theoretical Basis Document (ATBD) NASA Global Precipitation Measurement (GPM)
577 Integrated Multi-satellitE Retrievals for GPM (IMERG), Nasa.

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

580 Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.-L., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie,
581 P., 2019. Algorithm Theoretical Basis Document (ATBD) Version 06, NASA Global Precipitation
582 Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG).

583 Hussain, Y., Satgé, F., Hussain, M.B., Martinez-Caravajal, H., Bonnet, M.-P., Cardenas-Soto, M., Llacer Roig, H.,
584 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

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

Knoben, W.J.M., Freer, J.E., Fowler, K.J.A., Peel, M.C., Woods, R.A., 2019. Modular Assessment of Rainfall – Runoff Models Toolbox providing implementations of 46 conceptual hydrologic models as continuous state-space formulations. Geosci. Model Dev. 2463–2480. https://doi.org/https://doi.org/10.5194/gmd-12-2463-2019

Knoben, W.J.M., Freer, J.E., Peel, M.C., Fowler, K.J.A., Woods, R.A., 2020. A Brief Analysis of Conceptual Model Structure Uncertainty Using 36 Models and 559 Catchments. Water Resour. Res. 56, 1–23. https://doi.org/10.1029/2019WR025975

Levizzani, V., Amorati, R., Meneguzzo, F., 2002. A review of satellite-based rainfall estimation methods, … Commission Project MUSIC ….

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, 1101–1117. https://doi.org/10.1175/JHM-D-15-0190.1

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

Mazzoleni, M., Brandimarte, L., Amaranto, A., 2019. Evaluating precipitation datasets for large-scale distributed hydrological modelling. J. Hydrol. 578, 124076. https://doi.org/10.1016/j.jhydrol.2019.124076

Paiva, R.C.D., Buarque, D.C., Clarke, R.T., Collischonn, W., Allasia, D.G., 2011. Reduced precipitation over large water bodies in the Brazilian Amazon shown from TRMM data. Geophys. Res. Lett. 38, 2–6.

https://doi.org/10.1029/2010GL045277

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279, 275–289. https://doi.org/10.1016/S0022-1694(03)00225-7

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

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 Yuque Lima, W., 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., 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. https://doi.org/10.1016/j.jhydrol.2019.124412

Satgé, F., Hussain, Y., Bonnet, M.-P., Hussain, B., Martinez-Carvajal, H., Akhter, G., Uagoda, R., 2018. Benefits of the Successive GPM Based Satellite Precipitation Estimates IMERG–V03, –V04, –V05 and GSMaP–V06, –V07 Over Diverse Geomorphic and Meteorological Regions of Pakistan. Remote Sens. 10, 1373. https://doi.org/10.3390/rs10091373

Satgé, F, Hussain, Y., Molina-Carpio, J., Pillco, R., Laugner, C., Akhter, G., Bonnet, M.-P., 2020. Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. Int. J. Climatol. 41, E517–E536. https://doi.org/10.1002/joc.6704

Satgé, F., Hussain, Y., Xavier, A., Pillco, R., Salles, L., Timouk, F., Seyler, F., Garnier, J., Frappart, F., Bonnet, M., 2019a. Unraveling the impacts of droughts and agricultural intensi fi cation on the Altiplano water resources. Agric. For. Meteorol. 279, 107710. https://doi.org/10.1016/j.agrformet.2019.107710

Satgé, F., Ruelland, D., Bonnet, M., Molina, J., Pillco, R., 2019b. Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow- hydrological modelling in

637 the Lake Titicaca region. Hydrol. Earth Syst. Sci. 23, 595–619.
638 https://doi.org/https://doi.org/10.5194/hess-23-595-2019

639 Satgé, F., Xavier, A., Zolá, R., Hussain, Y., Timouk, F., Garnier, J., Bonnet, M.-P., 2017. Comparative Assessments
640 of the Latest GPM Mission's Spatially Enhanced Satellite Rainfall Products over the Main Bolivian
641 Watersheds. Remote Sens. 9, 369. https://doi.org/10.3390/rs9040369

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

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

648 Sun, Q., Miao, C., Duan, Q., Kong, D., 2015. Would the ' real ' observed dataset stand up ? A critical examination
649 of eight observed gridded climate datasets for China. Environ. Res. Lett. 9. https://doi.org/10.1088/1748-
650 9326/9/1/015001

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

654 Tarek, M., Brissette, F., Arsenault, R., 2019. Evaluation of the ERA5 reanalysis as a potential reference dataset
655 for hydrological modeling over North-America. Hydrol. Earth Syst. Sci. Discuss. 1–35.
656 https://doi.org/10.5194/hess-2019-316

657 Tarek, M., Brissette, F.P., Arsenault, R., 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset
658 for hydrological modelling over North America. Hydrol. Earth Syst. Sci. 24, 2527–2544.
659 https://doi.org/10.5194/hess-24-2527-2020

660 Tian, Y., Peters-Lidard, C.D., 2007. Systematic anomalies over inland water bodies in satellite-based
661 precipitation estimates. Geophys. Res. Lett. 34, L14403. https://doi.org/10.1029/2007GL030787

662 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, D24101. https://doi.org/10.1029/2009JD011949

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

Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H. V., Sorooshian, S., 2001. A framework for development and application of hydrological models. Hydrol. Earth Syst. Sci. 5, 13–26. https://doi.org/10.5194/hess-5-13-2001

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

Xavier, A.C., King, C.W., Scanlon, B.R., 2017. An update of Xavier, King and Scanlon (2016) daily precipitation gridded data set for the Brazil. An. do XVII Simp. Bras. Sensoriamento Remoto - SBSR 562–569.

Xavier, A.C., King, C.W., Scanlon, B.R., 2015. Daily gridded meteorological variables in Brazil (1980–2013). Int. J. Climatol. 36, 2644–2659. https://doi.org/10.1002/joc.4518

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., Lin, R., 2017. Reprocessed , Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998. J. hydrometeorol 18, 1617–1641. https://doi.org/10.1175/JHM-D-16-0168.1

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

Zhang, L., Ren, D., Nan, Z., Wang, W., Zhao, Yi, Zhao, Yanbo, Ma, Q., Wu, X., 2020. Interpolated or satellite-based precipitation? Implications for hydrological modeling in a meso-scale mountainous watershed on

689 the Qinghai-Tibet Plateau. J. Hydrol. 583, 124629. https://doi.org/10.1016/j.jhydrol.2020.124629

690

691 Adler, R.F., Huffman, G.J., Chang, A., Ferraro, F., Xie, P.-P., Janowiak, J., Rudolff, B., Schneider, U., Curtis, S.,
692 Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E., 2003. The Version-2 Global Precipitation
693 Climatology Project (GPCP) Monthly Precipitation Analysis (1979 – Present). J. Hydrometeorol. 1147–
694 1167.

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

698 Beck, H.E., Pan, M., Roy, T., Weedon, G.P., Pappenberger, F., Dijk, A.I.J.M. Van, Huffman, G.J., Adler, R.F.,
699 Wood, E.F., 2019. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the
700 CONUS. Hydrol. Earth Syst. Sci. 207–224. https://doi.org/https://doi.org/10.5194/hess-23-207-2019

701 Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A.I.J.M., Weedon, G., Brocca, L., Pappenberger, F.,
702 Huffman, G.J., Wood, E.F., 2017. Global-scale evaluation of 23 precipitation datasets using gauge
703 observations and hydrological modeling. Hydrol. Earth Syst. Sci. Discuss. 1–23.
704 https://doi.org/10.5194/hess-2017-508

705 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.
706 MSWEP V2 global 3-hourly 0.1° precipitation methodology and quantitative assessment. Bull. Am.
707 Meteorol. Soc. 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

708 Becker, A., Finger, P., Rudolf, B., Schamm, K., Schneider, U., Ziese, M., Precipitation, G., Centre, C.,
709 Wetterdienst, D., 2013. A description of the global land-surface precipitation data products of the Global
710 Precipitation Climatology Centre with sample applications including centennial (trend) analysis from
711 1901–present 71–99. https://doi.org/10.5194/essd-5-71-2013

712 Brocca, L., Massari, C., Pellarin, T., Filippucci, P., Ciabatta, L., Camici, S., Kerr, Y.H., Fernández-Prieto, D., 2020.
713 River flow prediction in data scarce regions: soil moisture integrated satellite rainfall products
714 outperform rain gauge observations in West Africa. Sci. Rep. 10, 1–14. https://doi.org/10.1038/s41598-

020-69343-x

Camici, S., Ciabatta, L., Massari, C., Brocca, L., 2018. How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean area. J. Hydrol. 563, 950–961. https://doi.org/10.1016/j.jhydrol.2018.06.067

Camici, S., Massari, C., Ciabatta, L., Marchesini, I., Brocca, L., 2020. Which rainfall score is more informative about the performance in river discharge simulation? A comprehensive assessment on 1318 basins over Europe. Hydrol. Earth Syst. Sci. 24, 4869–4885. https://doi.org/10.5194/hess-24-4869-2020

Chagas, V.B.P., L. B. Chaffe, P., Addor, N., M. Fan, F., S. Fleischmann, A., C. D. Paiva, R., Siqueira, V.A., 2020. CAMELS-BR: Hydrometeorological time series and landscape attributes for 897 catchments in Brazil. Earth Syst. Sci. Data 12, 2075–2096. https://doi.org/10.5194/essd-12-2075-2020

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

Dembélé, M., Hrachowitz, M., Savenije, H.H.G., Mariéthoz, G., Schaefli, B., 2020. Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets. Water Resour. Res. 56, 1–26. https://doi.org/10.1029/2019WR026085

Dembele, M., Schaefli, B., Van De Giesen, N., 2020. Suitability of 17 gridded rainfall and temperature datasets for large-scale hydrological modelling in West Africa. Hydrol. Earth Syst. Sci. 24, 5379–5406. https://doi.org/10.5194/hess-24-5379-2020

Fallah, A., O, S., Orth, R., 2020. Climate-dependent propagation of precipitation uncertainty into the water cycle. Hydrol. Earth Syst. Sci. Discuss. 1–19. https://doi.org/10.5194/hess-2019-660

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, 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, 4302–4315. https://doi.org/10.1002/joc.5086

741 Funk, Chris, Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L.,
742 Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations — a new
743 environmental record for monitoring extremes 1–21. https://doi.org/10.1038/sdata.2015.66

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

746 Gebregiorgis, A.S., Hossain, F., 2013. Understanding the Dependence of Satellite Rainfall Uncertainty on
747 Topography and Climate for Hydrologic Model Simulation. IEEE Trans. Geosci. Remote Sens. 51, 704–718.
748 https://doi.org/10.1109/TGRS.2012.2196282

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

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

757 Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic
758 observations - the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711

759 Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorroshian, S., Tan, J., Xie, P.,
760 2018. Algorithm Theoretical Basis Document (ATBD) NASA Global Precipitation Measurement (GPM)
761 Integrated Multi-satellitE Retrievals for GPM (IMERG), Nasa.

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

764 Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.-L., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie,
765 P., 2019. Algorithm Theoretical Basis Document (ATBD) Version 06, NASA Global Precipitation
766 Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG).

767 Hussain, Y., Satgé, F., Hussain, M.B., Martinez-Caravajal, H., Bonnet, M.-P., Cardenas-Soto, M., Llacer Roig, H.,
768 Akhter, G., 2017. Performance of CMORPH, TMPA and PERSIANN rainfall datasets over plain,
769 mountainous and glacial regions of Pakistan. Theor. Appl. Climatol. https://doi.org/10.1007/s00704-016-
770 2027-z

771 Jiang, L., Bauer-Gottwein, P., 2019. How do GPM IMERG precipitation estimates perform as hydrological model
772 forcing? Evaluation for 300 catchments across Mainland China. J. Hydrol. 572, 486–500.
773 https://doi.org/10.1016/j.jhydrol.2019.03.042

774 Knoben, W.J.M., Freer, J.E., Fowler, K.J.A., Peel, M.C., Woods, R.A., 2019. Modular Assessment of Rainfall –
775 Runoff Models Toolbox providing implementations of 46 conceptual hydrologic models as continuous
776 state-space formulations. Geosci. Model Dev. 2463–2480. https://doi.org/https://doi.org/10.5194/gmd-
777 12-2463-2019

778 Knoben, W.J.M., Freer, J.E., Peel, M.C., Fowler, K.J.A., Woods, R.A., 2020. A Brief Analysis of Conceptual Model
779 Structure Uncertainty Using 36 Models and 559 Catchments. Water Resour. Res. 56, 1–23.
780 https://doi.org/10.1029/2019WR025975

781 Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E., 1998. Convergence properties of the nelder–mead simplex
782 method in low dimensions ∗. SIAM J. Optim. 9.
783 https://doi.org/https://doi.org/10.1137/S1052623496303470

784 Levizzani, V., Amorati, R., Meneguzzo, F., 2002. A review of satellite-based rainfall estimation methods, …
785 Commission Project MUSIC ….

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

788 Maggioni, V., Meyers, P.C., Robinson, M.D., 2016. A Review of Merged High-Resolution Satellite Precipitation
789 Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. J. Hydrometeorol. 17,
790 1101–1117. https://doi.org/10.1175/JHM-D-15-0190.1

791 Martens, B., Miralles, D.G., Lievens, H., Van Der Schalie, R., De Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E.,
792 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

Mazzoleni, M., Brandimarte, L., Amaranto, A., 2019. Evaluating precipitation datasets for large-scale distributed hydrological modelling. J. Hydrol. 578, 124076. https://doi.org/10.1016/j.jhydrol.2019.124076

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279, 275–289. https://doi.org/10.1016/S0022-1694(03)00225-7

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

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., 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. https://doi.org/10.1016/j.jhydrol.2019.124412

Satgé, F., Hussain, Y., Bonnet, M.-P., Hussain, B., Martinez-Carvajal, H., Akhter, G., Uagoda, R., 2018. Benefits of the Successive GPM Based Satellite Precipitation Estimates IMERG–V03, –V04, –V05 and GSMaP–V06, –V07 Over Diverse Geomorphic and Meteorological Regions of Pakistan. Remote Sens. 10, 1373. https://doi.org/10.3390/rs10091373

Satgé, F, Hussain, Y., Molina-Carpio, J., Pillco, R., Laugner, C., Akhter, G., Bonnet, M.-P., 2020. Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. Int. J. Climatol. 41, E517–E536. https://doi.org/10.1002/joc.6704

Satgé, F., Hussain, Y., Xavier, A., Pillco, R., Salles, L., Timouk, F., Seyler, F., Garnier, J., Frappart, F., Bonnet, M., 2019a. Unraveling the impacts of droughts and agricultural intensi fi cation on the Altiplano water resources. Agric. For. Meteorol. 279, 107710. https://doi.org/10.1016/j.agrformet.2019.107710

Satgé, F., Ruelland, D., Bonnet, M., Molina, J., Pillco, R., 2019b. 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/https://doi.org/10.5194/hess-23-595-2019

Satgé, F., Xavier, A., Zolá, R., Hussain, Y., Timouk, F., Garnier, J., Bonnet, M.-P., 2017. Comparative Assessments of the Latest GPM Mission's Spatially Enhanced Satellite Rainfall Products over the Main Bolivian Watersheds. Remote Sens. 9, 369. https://doi.org/10.3390/rs9040369

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. https://doi.org/10.1007/s00704-013-0860-x

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

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

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

Tarek, M., Brissette, F., Arsenault, R., 2019. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modeling over North-America. Hydrol. Earth Syst. Sci. Discuss. 1–35. https://doi.org/10.5194/hess-2019-316

Tarek, M., Brissette, F.P., Arsenault, R., 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrol. Earth Syst. Sci. 24, 2527–2544. https://doi.org/10.5194/hess-24-2527-2020

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, D24101. https://doi.org/10.1029/2009JD011949

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

Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H. V., Sorooshian, S., 2001. A framework for development and application of hydrological models. Hydrol. Earth Syst. Sci. 5, 13–26. https://doi.org/10.5194/hess-5-13-2001

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

Xavier, A.C., King, C.W., Scanlon, B.R., 2017. An update of Xavier, King and Scanlon (2016) daily precipitation gridded data set for the Brazil. An. do XVII Simp. Bras. Sensoriamento Remoto - SBSR 562–569.

Xavier, A.C., King, C.W., Scanlon, B.R., 2015. Daily gridded meteorological variables in Brazil (1980–2013). Int. J. Climatol. 36, 2644–2659. https://doi.org/10.1002/joc.4518

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., Lin, R., 2017. Reprocessed , Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998. J. hydrometeorol 18, 1617–1641. https://doi.org/10.1175/JHM-D-16-0168.1

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

Zhang, L., Ren, D., Nan, Z., Wang, W., Zhao, Yi, Zhao, Yanbo, Ma, Q., Wu, X., 2020. Interpolated or satellite-based precipitation? Implications for hydrological modeling in a meso-scale mountainous watershed on the Qinghai-Tibet Plateau. J. Hydrol. 583, 124629. https://doi.org/10.1016/j.jhydrol.2020.124629