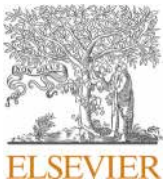
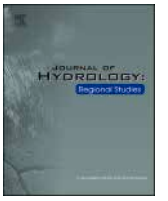
Journal of Hydrology: Regional Studies 47 (2023) 101400



Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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



Ensemble precipitation estimates based on an assessment of 21 gridded precipitation datasets to improve precipitation estimations across Madagascar

Camille C. Ollivier [a], Simon D. Carrière [b,*], Thomas Heath [c], Albert Olioso [d], Zo Rabefitia [e], Heritiana Rakoto [f], Ludovic Oudin [b], Frédéric Satgé [g]

[a] Independent researcher, Grenoble, France
[b] Sorbonne Université, UPMC, CNRS, EPHE, UMR 7619 METIS, 4 place Jussieu, 75005 Paris, France
[c] Action Contre la Faim, Paris, France
[d] Unité Forêt Méditerranéenne, INRAE, Avignon, France
[e] Direction Générale de la Météorologie, Antananarivo, Madagascar
[f] Institut et Observatoire de Géophysique d'Antananarivo, Université d'Antananarivo, Madagascar
[g] IRD, University of Montpellier, Espace-DEV, Montpellier, France

ARTICLE INFO

Keywords:
Precipitation products
Remote sensing
Ensemble approach
Hydrology
Madagascar

ABSTRACT

Study region: this study focuses on Madagascar. This island is characterized by a great diversity of climate, due to trade winds and the varying topography. This country is also undergoing extreme rainfall events such as droughts and cyclones.
Study focus: the rain gauge network of Madagascar is limited (about 30 stations). Consequently, we consider relevant satellite-based precipitation datasets to fill gaps in ground-based datasets. We assessed the reliability of 21 satellite-based and reanalysis precipitation products (P-datasets) through a direct comparison with 24 rain gauge station measurements at the monthly time step, using four statistical indicators: Kling-Gupta Efficiency (KGE), Correlation Coefficient (CC), Root Mean Square Error (RMSE), and Bias. Based on this first analysis, we produced a merged dataset based on a weighted average of the 21 products.
New hydrological insights for the region: based on the KGE and the CC scores, WFDEI (WATCH Forcing Data methodology applied to ERA-Interim), CMORPH-BLD (Climate Prediction Center MORPHing satellite-gauge merged) and MSWEP (Multi-Source Weighted Ensemble Precipitation) are the most accurate for estimating rainfall at the national scale. Additionally, the results reveal a high discrepancy between bio-climatic regions. The merged dataset reveals higher performance than the other products in all situations. These results demonstrate the usefulness of a merging approach in an area with a deficit of rainfall data and a climatic and topographic diversity.

* Corresponding author.
E-mail address: simon.carriere@upmc.fr (S.D. Carrière).

https://doi.org/10.1016/j.ejrh.2023.101400
Received 27 December 2022; Received in revised form 24 April 2023; Accepted 26 April 2023
Available online 28 April 2023

1. Introduction

1.1. Precipitation: a key variable worldwide

Precipitation is the primary process for redistributing water on the Earth. Precipitation monitoring provides vital information for society in the areas of flood monitoring, drought, agricultural production, and public health (World Health Organization, 2015). Rain gauge networks are the main source of precipitation measurement data. Rain gauge networks provide point scale measurements, whereas the spatial distribution of precipitation is highly heterogeneous, from the local scale (Lebel et al., 2009) to the regional scale (Depraetere et al., 2009). Rain gauge networks are rarely dense enough for scientists to be able to estimate the spatial heterogeneity of precipitation (Sun et al., 2018). Rain gauge data can be combined with weather radar to provide robust precipitation mapping (Collier, 1986; Sinclair and Pegram, 2005). However, combination of instrumentation is rarely available in developing countries where weather monitoring receives scant investment (Kidd and Huffman, 2011).

1.2. Gridded precipitation dataset opportunity

The development of quasi-global gridded precipitation datasets based on satellite or reanalysis data (P-datasets) with regular spatial and temporal coverage offers a promising alternative to traditional gauge measurements. This is especially true for remote regions where the socio-economic context prevents the maintenance of a dense station network in space and time.

P-datasets are built from three types of data (i.e. satellites, reanalysis, and rain gauges) (Gehne et al., 2016; Satgé et al., 2020). Microwave and/or infrared sensors attached to satellites provide the satellite data used to estimate precipitation (Huffman, 2020). Reanalysis data come from physical and dynamical numerical model data (Poccard-leclercq, 2000). Rain gauge data are interpolated to produce spatially continuous data. Most P-datasets combine multiple sources of data (i.e. satellites, reanalysis, and rain gauges) to improve their precipitation estimates. Over the past 25 years, the successive release of P-datasets that incorporate technical advances has led to improvement (Nwachukwu et al., 2020). The National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) launched the Tropical Rainfall Measuring Mission (TRMM) in 1997, making it possible to develop P-datasets that could estimate precipitation at the 0.25° grid scale, such as the TRMM multi-satellite precipitation analysis dataset (TMPA), the Climate Prediction Centre Morphing dataset (CMORPH), and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks dataset (PERSIANN). The Global Precipitation Mission (GPM) was launched in 2014 to overcome technical difficulties arising from the TRMM's aging mechanisms and sensors, and to ensure continued precipitation data acquisition. With the GPM mission, several P-datasets that can estimate precipitation at the 0.1° grid scale have been developed, such as the Integrated Multi-satellite Retrievals for the GPM dataset (IMERG) and the Global Satellite Mapping of Precipitation dataset (GSMaP). Finally, several P-datasets have been developed using previous P-datasets and missions to estimate precipitation over a larger time window, such as the PERSIANN-Climate Data Record (PERSIANN-CDR), the Multi-Source Weighted-Ensemble Precipitation (MSWEP), the Climate Hazards Group Infrared Precipitation (CHIRP) with station data (CHIRPS), the Tropical Applications of Meteorology using SATellite and ground-based observation (TAMSAT), and the African Rainfall Climatology (ARC v.2) (Nwachukwu et al., 2020).

1.3. Gridded precipitation spatial uncertainties

All P-datasets are subject to error. Uncertainties arise from the techniques and methodologies used, and their performance is also affected by background conditions, the season, latitude, and meteorological conditions (Kidd and Huffman, 2011). Sensor deficiencies can also be a source of error (Rahman et al., 2020). Due to these various constraints, no P-dataset has been validated as suitable worldwide application (Gehne et al., 2016). Therefore, they need to be evaluated before being used for a specific purpose, such as regional climate and hydrological studies.

One way to assess P-dataset reliability is to compare its estimate with the measurement at the rain gauge location. A study assessing the reliability of seven P-datasets by direct comparison with rain gauge observations across the South American Andean Plateau shows that the P-dataset performance varies with the latitude, the longitude, and the presence of water surfaces (Satgé et al., 2016). In Pakistan, Rahman et al. (2020) observed strong variability of P-datasets performances according to the orographic and climatic conditions. This makes the results of these studies not easily transferable to other climates and other regions.

In West Africa, the reliability of 23 P-datasets was assessed through direct comparison with rain gauge observations collected from several African national meteorological and hydrological centers (Satgé et al., 2020). Overall, the most satisfactory P-datasets were CHIRPS v.2 and MSWEP v.2.2. However, the best performing P-dataset showed wide variation in space at the gauge scale. The authors suggest that to obtain the most accurate monthly precipitation representation across West Africa, ideally 14 out of the 23 P-datasets analyzed should be considered. In Nigeria, Nwachukwu and colleagues used data from 11 rain gauge stations to evaluate 16 P-datasets, using rain gauge stations that were widely distributed across the country (Nwachukwu et al., 2020). The results show that P-dataset accuracy depends on the region. In fact, the authors conclude the study by suggesting that to obtain the most accurate monthly precipitation representation across Nigeria, 5 out of the 16 P-datasets analyzed should be considered.

1.4. Toward enhanced gridded precipitation datasets?

To take advantage of P-dataset performance spatial variability, several studies highlight the benefits of applying ensemble-merging

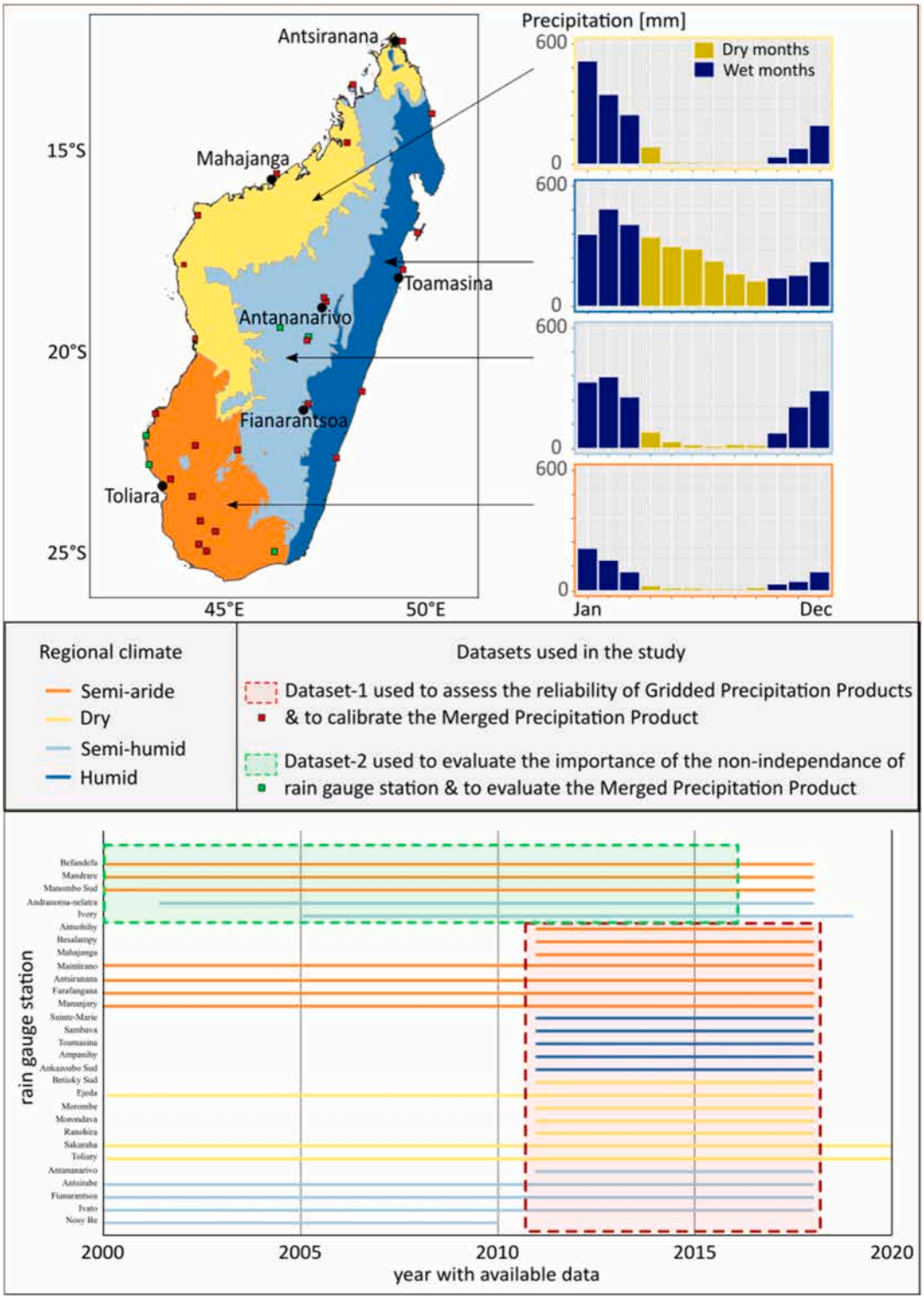


Fig. 1. The diagrams correspond to the average monthly precipitation totals based on all rain gauges for each climatic region for the period 2011–2018. Bottom: period with available data for each rain gauge.
Top: Location map of Malagasy rain gauges and the four climatic regions (adapted from Cornet, 1974).

approaches to multiple P-datasets to improve precipitation estimates (Grimes, 2008). These approaches are, for example, implemented in the creation of P-datasets combining several sources of data, like MSWEP. However, these P-datasets are optimized at a global scale, and not necessarily adequate at the local scale.

Consequently, several studies are developing a local ensemble-merging approach. In Japan, three merging techniques were tested for four P-datasets: average, error variance, and inverse error variance weighting (Mastrantonas et al., 2019). The results show that all merging techniques outperformed individual P-dataset estimations. In Pakistan, Rahman et al. (2018) tested several methods, and the results displayed spatial variability in performance. The merged product outperformed IMERG and TMPA on plains and moderately elevated areas but not in highly elevated areas. Therefore, Rahman et al. (2021) implemented a weighting method that varies both spatially and temporally (Weighted Average Least Squares (WALS)). The spatial and temporal weighting variation greatly reduced the uncertainties associated with an individual P-dataset.

Overall, these studies show that the ensemble-merging approach applied to multiple P-datasets leads to more accurate precipitation estimates than estimates derived from single P-datasets.

1.5. Study objectives

Given the previously described context, the two main objectives of this study are as follows:

- Assess for the first time the reliability of 21 P-datasets in Madagascar at the monthly time step. This first step will complement a previous study led by Randriatsara et al. (2022), who assessed 15 P-datasets at annual and seasonal time steps, and one study conducted by Ramahaimandimby et al. (2022) who evaluated six P-datasets at the daily time step over one watershed.
- Propose an ensemble-merging method using the most efficient P-datasets to improve national precipitation monitoring and leading to a model that could be applied to other regions of the world.

Madagascar is chosen as the study area because of its climatic diversity, which is expected to affect P-dataset reliability spatially, and as a result, introduces the need for the ensemble-merging method. Moreover, the low density of the current rain gauge network pushes water and climate science toward the use of these P-datasets. Therefore, this study is a step toward improved precipitation monitoring and related studies across Madagascar.

2. Materials

2.1. Study area

Madagascar, one of the world's largest islands, has a total area of 587,041 km^2. Located in the Indian Ocean, Madagascar is characterized by topographic diversity, with high plateaus at its center, a relief of plains and lower plateaus toward the west, and a narrow coastal strip in the east. The highest peak, Maromokotro has an elevation of 2876 m. This topography induces a strong heterogeneity in precipitation distribution. The eastern part, which is exposed to trade winds, experiences more rain ($\sim$ 3500 mm $year^{-1}$) than the western part ($\sim$ 800 mm $year^{-1}$). Four bioclimatic regions (i.e. humid, semi-humid, dry, and semi-arid) have been identified on the island (Cornet, 1974, Fig. 1) and two distinct seasons: wet (November to April) and dry (May to October). This seasonality is more pronounced in the semi-arid region than in the humid region (diagrams in Fig. 1). Madagascar is also regularly exposed to cyclones (Aruna 2013, Fundi 2015, Eketsang 2019) that are responsible for the high interannual variability in precipitation.

Madagascar is a classic example of a land mass where it is crucial but difficult to characterize the water cycle. The country has one of the highest global poverty rates (World Bank, 2020) and the southwestern district, a semi-arid area, is the poorest region in the country. This poverty is associated with malnutrition and health problems due to drought (UNICEF, 2018, 2019). Drought often triggers humanitarian crises that involve access to water (hygiene and drinking water) and food (development of crops and livestock), (FEWSNET, 2022). Understanding the water cycle is crucial for anticipating humanitarian crises. At the same time, Madagascar is a developing country where local institutions have difficulty maintaining a hydrological monitoring network. As a result, Madagascar has a low-density network of meteorological stations. Since 2014, the GROSoM observatory (Groundwater Resource Observatory for Southwestern Madagascar) has been producing meteorological data to fill this data gap in the semi-arid region (Carrière et al., 2021).

2.2. Rain gauges

Twenty-nine meteorological station datasets were obtained from the Malagasy National Meteorological Service (DGM – Direction Générale de la Météorologie), the French Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), and from Non-Governmental Organizations (NGOs).

The 29 meteorological station datasets were split in two groups. The first group (Dataset-1: red squares in Fig. 1) includes meteorological stations (24) that belong to the Malagasy Meteorological Directorate. The second group (Dataset-2: green squares in Fig. 1) consists of meteorological stations (5) that belong to other institutions (NGOs, Cooperative). Because Dataset-2 includes only stations that are not part of the Malagasy Meteorological Directorate network and are not referenced in global rain gauge datasets (see Appendix Table A.2), this dataset is considered to be independent from P-datasets that use gauge information for post calibration.

It should be noted that some stations have daily records whereas others have only monthly records. To take full advantage of the existing rain gauge network, the analysis is conducted at the monthly time step. For daily rain gauges, monthly amounts were

Table 1
Main characteristics of the gridded precipitation datasets. In the data source column, **S, R,** and **G** stand for satellite, reanalysis, and rain gauge data. Spatial coverage refers to the absolute maximum and minimum latitudes with precipitation information. Latency refers to the time delay for data availability. Res. = Resolution. Cov. = Coverage.

Name	Full name	Data	Spatial res.	Spatial cov.	Temporal res.	Temporal cov.	Latency
ARC-2	African Rainfall Climatology	S, G	0.1°	Africa	Day	1983 - present	2 days
CHIRP v.2	Climate Hazards Group InfraRed	S, R	0.05°	50°	Day	1981 - present	2 days
CHIRPS v.2	CHIRP with station	S, R, G	0.05°	50°	Day	1981 - present	1 month
CMORPH-BLD v.1	Climate Prediction Center MORPHing satellite- gauge merged	S, G	0.25°	60°	Day	1998 - present	18 h
CMORPH-CRT v.1	Climate Prediction Center MORPHing bias corrected	S, G	0.25°	60°	3 h	1998 - present	18 h
ERA-5 Land	European Centre for Medium-range Weather Forecast Re Analysis	R	0.25°	Global-Land	1 h	1979 - present	5 days
GSMaP-Adj v.6	Global Satellite Mapping of Precipitation adjusted	S, G	0.1°	60°	1 h	2000 - present	4 h
GSMaP-RT v.6	Global Satellite Mapping of Precipitation standard	S	0.1°	60°	1 h	2000 - present	4 h
IMERG_ER v.6	Integrated Multi-Satellite Retrievals for GPM Early Run (Full coverage)	S	0.1°	60°	0.5 h	2000 - present	4 h
IMERG_FR v.6	Integrated Multi-Satellite Retrievals for GPM Final Run (Full coverage)	S, G	0.1°	60°	0.5 h	2000 - present	3.5 month
IMERG_LR v.6	Integrated Multi-Satellite Retrievals for GPM Late Run (Full coverage)	S	0.1°	60°	0.5 h	2000 - present	½ day
MERRA2_LND	Modern-Era Retrospective analysis for Research and Applications - Land	S, R, G	0.625°× 0.5°	Global-Land	1 h	1980 - present	3 weeks
MERRA2-FLX	Modern-Era Retrospective analysis for Research and Applications - Fluxes	S, R, G	0.625°× 0.5°	Global	1 h	1980 - present	3 weeks
MSWEP v.2.2	Multi-Source Weighted Ensemble Precipitation	S, R, G	0.1°	Global	3 h	1979 - present	3 h
PRISM	Precipitation Inferred from Soil Moisture	S	0.1°	Africa	3 h	2010 - present	1 day
PERSIANN-CDR	Precipitation Estimates from Remotely Sensed Information using Artificial Neural Network and Climate Data Record	S, G	0.25°	60°	Day	1983 - present	3 months
TAMSAT v.3	Tropical Applications of Meteorology using SATellite and ground-based observations	S, G	0.0375°	Africa	Day	1983 - present	2 days
TMPA-Adj v.7	TRMM Multi-satellite Precipitation Analysis Real Time Adjusted	S, G	0.25°	50°	3 h	1998–2019	Stopped
TMPA-RT v.7	TRMM Multi- satellite Precipitation Analysis Real Time	S	0.25°	60°	3 h	2000–2019	Stopped
WFDEI_CRU	WATCH Forcing Data methodology applied to ERA-Interim using CRU TS3.101/TS3.21 precipitation totals	R, G	0.5°	Global-Land	3 h	1979–2018	Stopped
WFDEI_GPCC	WATCH Forcing Data methodology applied to ERA-Interim using GPCCv5/v6 precipitation totals	R, G	0.5°	Global-Land	3 h	1979–2016	Stopped

calculated only for months that have 95% or more of the daily data. Despite some missing months, Dataset-1 consists of a continuous observation period of 6 years (2011–2016), whereas Dataset-2 contains a continuous 17-year observation period (2000–2016).

2.3. Gridded precipitation datasets

We selected twenty-one satellite-based and reanalysis P-datasets (Table 1). GSMaP-RT v.6, IMERG-ER v.6, IMERG-LR v.6, PRISM and TMPA-RT v.7 are based on satellite data (**S**). ERA-5 Land uses reanalysis data (**R**). CHIRP v.2 combines both **S** and **R** data. WDFEI-CRU and WDFEI-GPCC combine **R** and **G** data. Eight P-datasets combine **S** and **G** data: ARC-2, CMORPH-BLD v.1, CMORPH-CRT v.1, GSMaP-Adj v.6, IMERG-FR v.6, PERSIANN-CDR, TAMSAT v.3, TMPA-Adj v.7. Finally, four P-datasets combine all three data sources **S, R, G:** CHIRPS v.2, MERRA2-LND, MERRA2-FLX, MSWEP v.2.2.

Fig. 2 shows the mean annual precipitation retrieved from the 21 P-datasets for the period 2007–2016 (2010–2021 for PRISM). For most datasets, the estimated annual precipitation patterns are consistent with the climatic regions. However, the P-datasets have contrasting total estimated precipitation amounts. According to ARC-2, annual precipitation remains below 1500 mm year^{-1} over the entire island, whereas according to MERRA2-FLX, precipitation reaches 3000 mm year^{-1} in the humid region. These differences support the need for an evaluation study in this area.

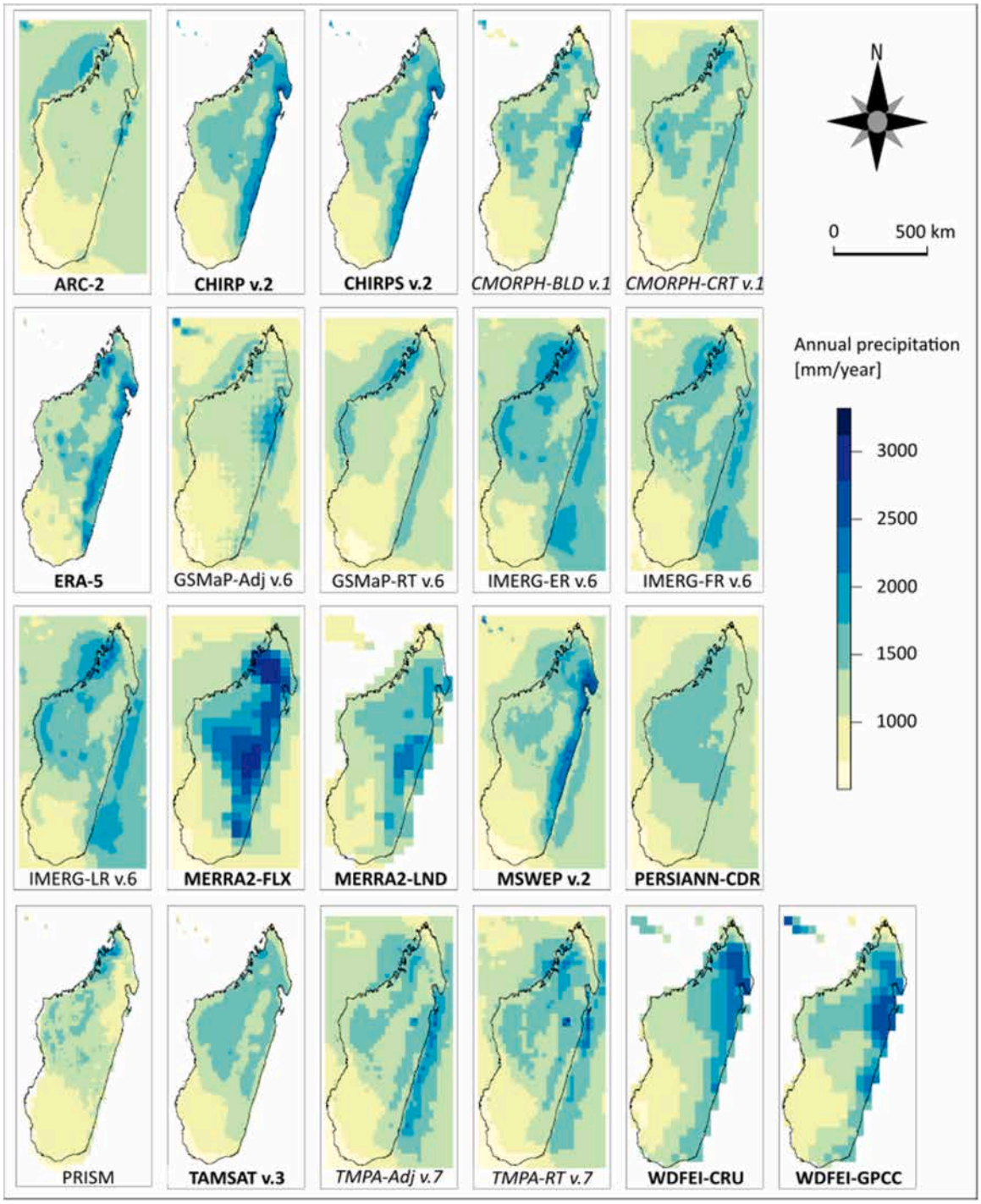


Fig. 2. Annual precipitation totals in Madagascar according to the estimates of twenty-one gridded precipitation datasets (average values in mm/year for the period 2007–2016, and for the period 2011–2016 for the product PRISM). P-datasets in *italics* are based on the TRMM mission, P-datasets in regular font are based on the GPM mission. Long-term P-dataset are in bold.

3. Methodology

First, we assess the reliability of P-datasets for Madagascar. Then we develop a new Merged Product Dataset. The major steps, explained thereafter, are summarized in Fig. 3.

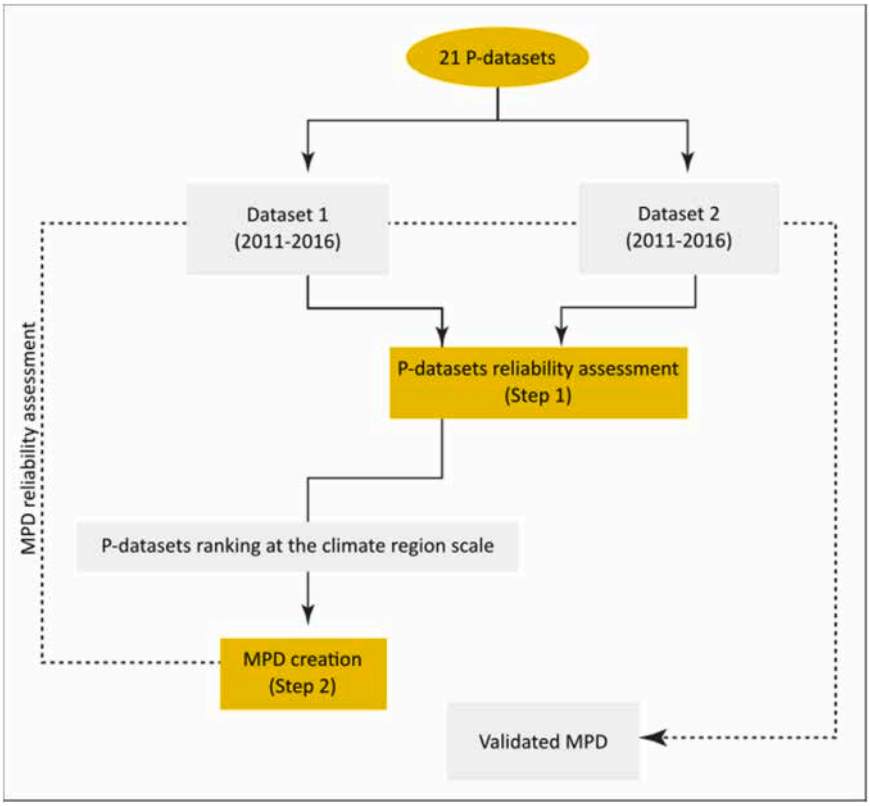


Fig. 3. Flow chart of the methodology implemented for the P-datasets reliability assessment and the MPD creation.

3.1. Gridded precipitation dataset - pre-processing

The method we used is based on a point-to-pixel comparison between P-datasets and rain gauge measurements. Precipitation data from the P-dataset grid-cells that match rain gauge locations were extracted. Because the P-datasets use different spatial resolutions, some were rescaled. Most have a resolution of 0.1° (see Table 1). P-datasets with a spatial resolution of less than 0.1° were resampled to 0.1° by averaging, and those with spatial resolution greater than 0.1° were resampled to 0.1° using bilinear interpolation. Temporal resolution ranged from 30 min to one day (see Table 1). All P-dataset data were aggregated to a monthly time step. Months common to P-datasets and rain gauge datasets were extracted once from all datasets.

3.2. Gridded precipitation dataset - reliability assessment

The regridded P-datasets were compared to rain gauge data using four statistical indicators: Root Mean Square Error (RMSE), Bias, Correlation Coefficient (CC) and Kling-Gupta Efficiency (KGE). RMSE (Eq. 1) provides the magnitude of the mean error; its optimum value is 0. Bias (Eq. 2) indicates the tendency of the P-dataset to overestimate or underestimate the precipitation amount; its optimum value is 0. The Pearson linear correlation coefficient (Eq. 3) indicates how well a P-dataset reproduces temporal precipitation variability; its optimum value is 1. The correlation coefficient is then used to calculate KGE (Eq. 4), which indicates the P-dataset's ability to reproduce temporal dynamics and to estimate means (Gupta et al., 2009; Kling et al., 2012); its optimum value is 1.

$$RMSE = \sqrt{\frac{1}{n}\sum(v_o - v_s)^2} \tag{1}$$

$$Bias = \frac{1}{n}\sum(v_s - v_o) \tag{2}$$

$$CC = \frac{1}{n}\sum_{i=1}^{n} \frac{(v_o - \mu_{vo})(v_s - \mu_{vs})}{\sigma_{vo} * \sigma_{vs}} \tag{3}$$

$$KGE = 1 - \sqrt{(CC - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{4}$$

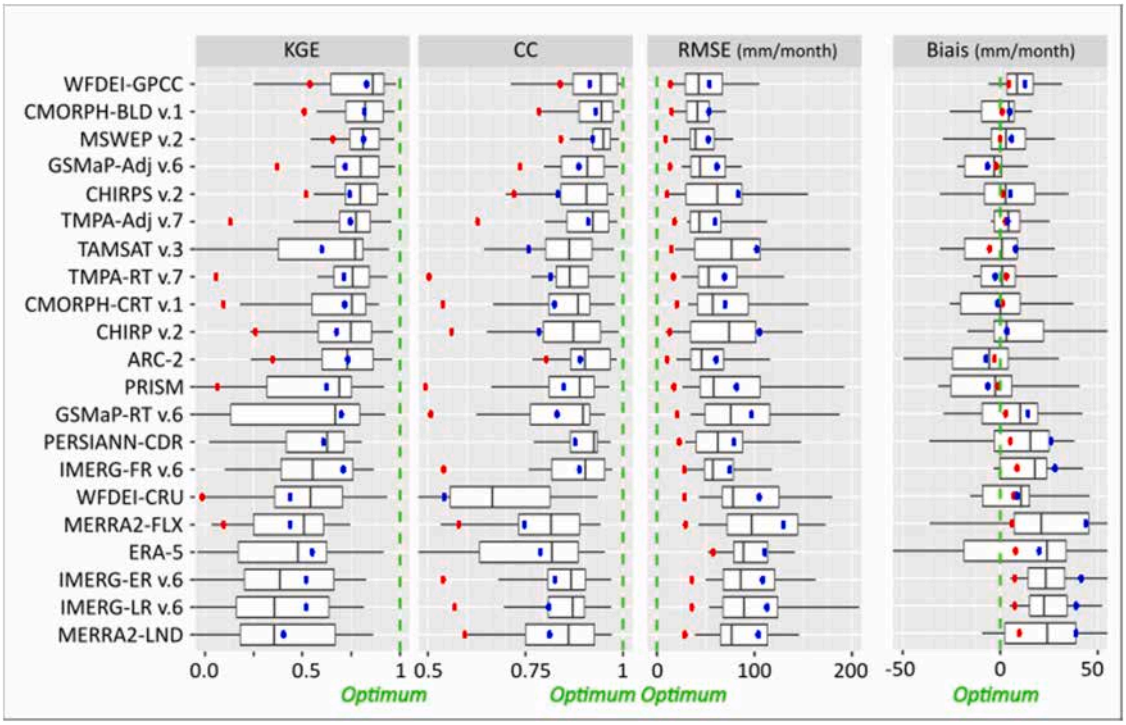


Fig. 4. Reliability assessment for 21 P-datasets at the national level, using monthly data from 24 meteorological stations (period 2011–2016). The dotted green line is the optimum for each indicator. The right and left sides of the boxes represent the 25th and 75th percentile values, respectively. The blue dot is the median value for the wet season, and the red dot is the median value for the dry season. The P-datasets are ranked by the median value of KGE.

with $\beta = \frac{\frac{\sigma_{vs}}{\mu_{vs}}}{\frac{\sigma_{vo}}{\mu_{vo}}}$ and $\gamma = \frac{\mu_{vs}}{\mu_{vo}}$ With n = number of measurements, v_0 = rain gauge value, v_s P-dataset value, σ = standard deviation, μ = mean.

For each rain gauge, the four statistical indicators were calculated using (i) all available months, (ii) the months corresponding to the wet season, i.e., from November to April, and (iii) the months corresponding to the dry season, i.e., from May to October. The analysis was initially performed at the national scale and then at the regional scale for the four climatic regions depicted in Fig. 1.

The P-datasets were first evaluated using Dataset-1 for the period 2011–2016. However, some Dataset-1 rain gauges may have been used in calculating some of the P-datasets. This non-independence of rain gauge data could result in an overestimate of the P-dataset capacity. Therefore, we repeated the analysis with independent rain gauges (Dataset-2), using the same statistical indicators for the period 2000–2016.

3.3. Gridded precipitation datasets - merging approach

Based on the P-dataset reliability assessment (Dataset-1), P-dataset reliability is expected to vary spatially depending on the region under consideration. As a result, we used a learning process incorporating the ensemble theory to create a merged gridded precipitation dataset (MPD).

First, we computed the median KGE value obtained from all stations ($KGE_{MedProd}$) for each P-dataset and for the entire climatic region under consideration. Then, for each region, the P-datasets were ranked from best to worst according to their $KGE_{MedProd}$.

Next, for each climatic region and P-dataset, the $KGE_{MedProd}$ was normalized (KGE_{norm}) by a Min-Max normalization (Eq. 5, where KGE_{min} = Minimum of the $KGE_{MedProd}$, and KGE_{max} = Maximum of the $KGE_{MedProd}$). The weighting of each dataset (w_{KGE}) is the ratio between KGE_{norm} and $\sum KGE_{norm}$ (Eq. 6). Consequently, the sum of the weighted KGEs equals 1. Because weightings were calculated separately for each climatic region and for a common period (2011–2016), the weightings vary according to location (climatic region) but not over time.

$$KGE_{norm(i)} = \frac{KGE_{MedProd}(i) - KGE_{\min}}{(KGE_{\max} - KGE_{\min})} \tag{5}$$

$$w_{KGE(i)} = \frac{KGE_{norm(i)}}{\sum KGE_{norm}} \tag{6}$$

Where i is the P-dataset number.

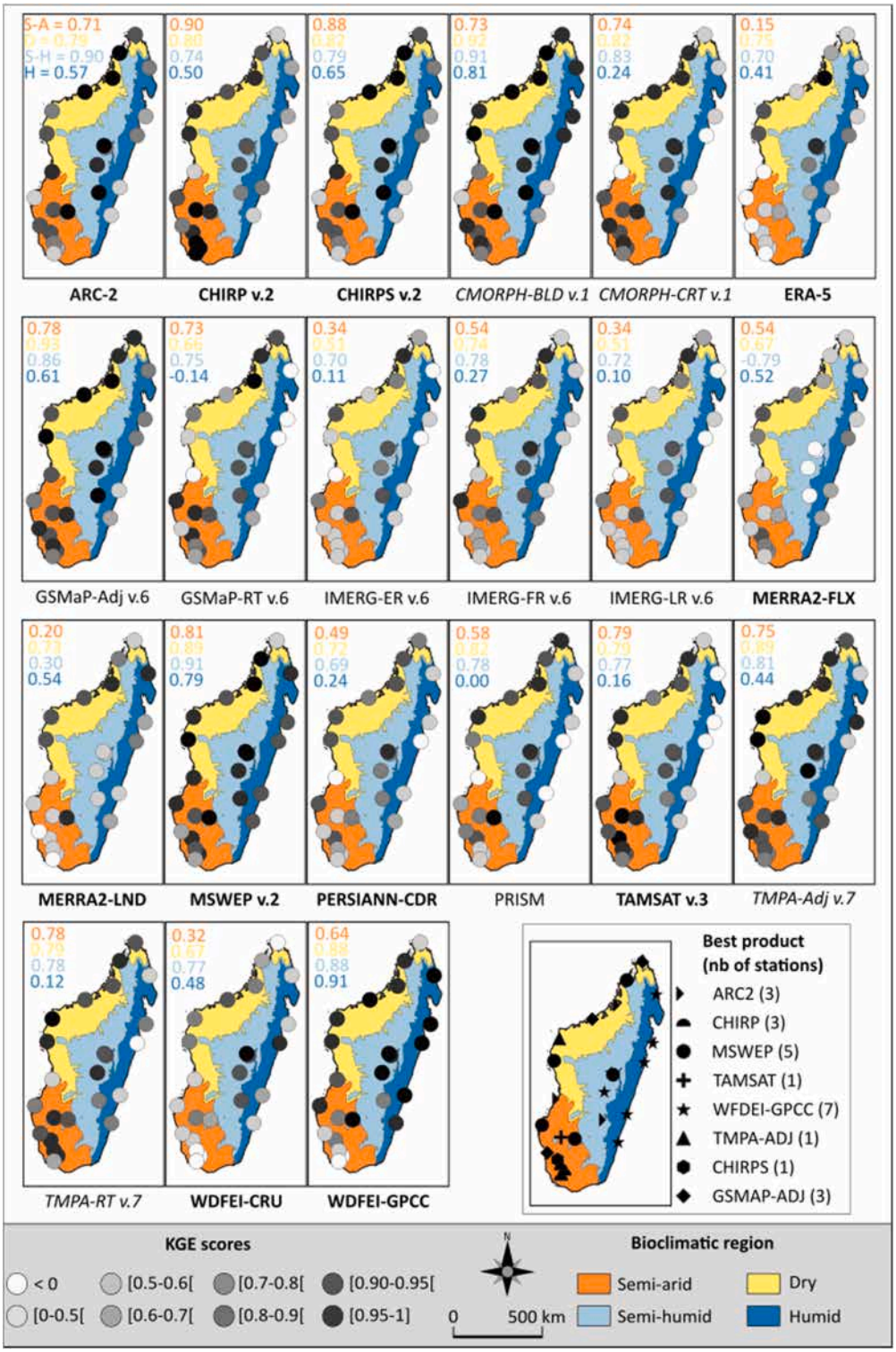


Fig. 5. KGE scores obtained with monthly data for the period 2011–2016 are shown per station for the 21 P-datasets. KGE scores according to the climatic region are shown in the upper left corner. The best product map indicates the product with the highest KGE for each rain gauge station.

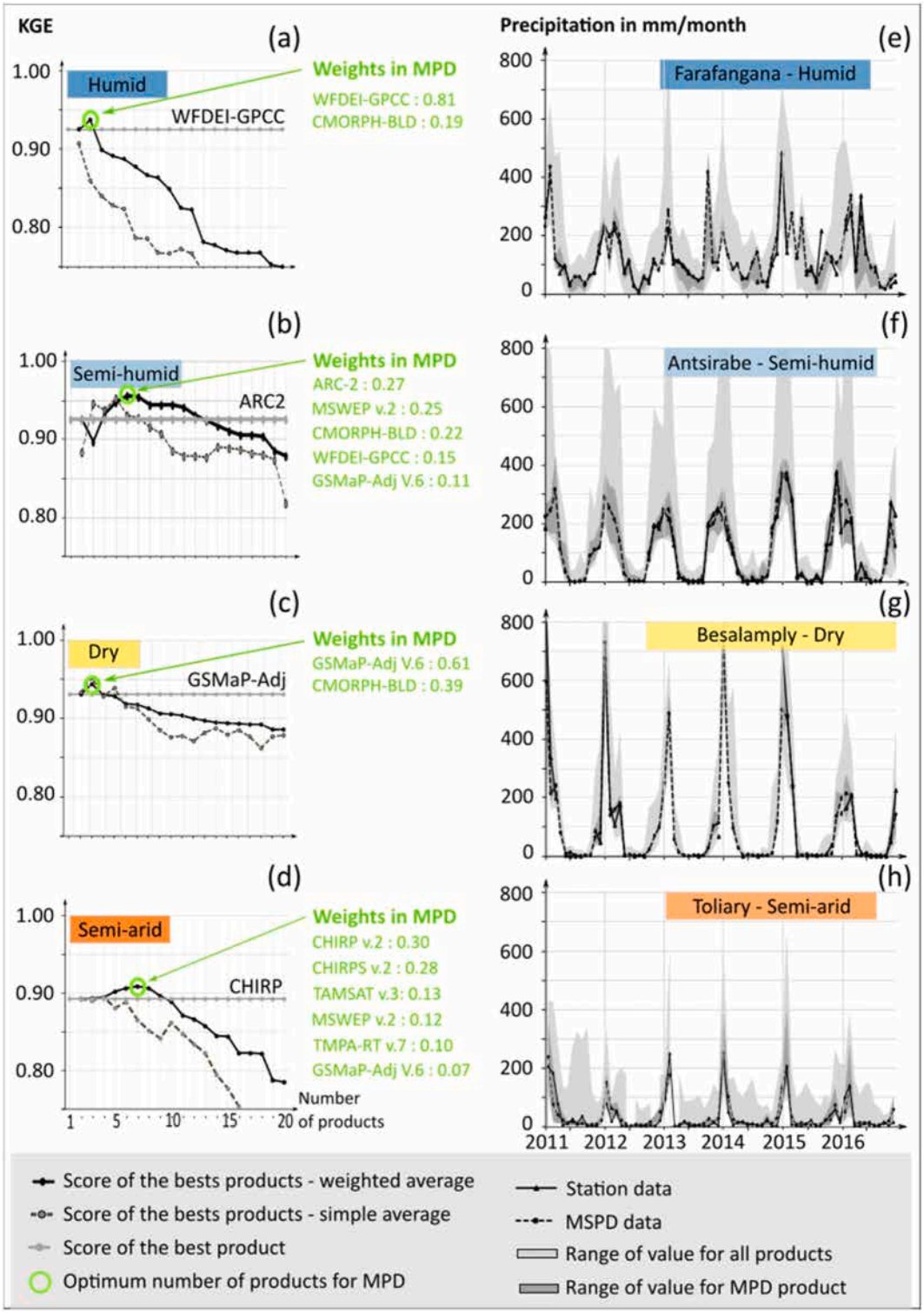


Fig. 6. (a), (b), (c), (d): for each climatic region, KGE as a function of the number of products (from 2 to 21 products ranked in descending order of the median KGE) for the weighted average (solid black line) and the simple median (dotted gray line). The best product is the gray line. Green circle indicates the optimum number of P-datasets. (e), (f), (g), (h): for 4 stations representing the four climatic regions, monthly precipitation totals between 2011 and 2016. Light gray = minimum and maximum values of the 21 products. Dark gray = minimum and maximum values of the MPD products. Solid line = MPD values. Dotted line = rain gauge values. For better readability of the graph, the maximum estimated precipitation is not shown when it is higher than 800 mm/month.

Finally, an arithmetic average and a weighted average are tested to create the MPD dataset. For each climatic region, the averaging processes (arithmetic and weighted) are first done based on the two most reliable P-datasets (i.e. descending $KGE_{MedProd}$). Then the averaging processes are reiterated, adding the next most reliable P-datasets one by one. By doing so, 20 combinations are considered in the averaging process to develop 20 MPD product versions (i.e by considering 2–21 P-datasets, see Fig. 3). The most accurate version is selected from these 20 versions. We used the term MPD_{wa} and MPD_{aa} to denote the MPD using weighted average and arithmetic average, respectively.

To assess the temporal transferability of the proposed ensemble-merging approach, MPD weightings were calibrated with Dataset-1 (period 2011–2016) and then evaluated with Dataset-2 (period 2000–2016).

4. Results

4.1. P-dataset reliability assessment

4.1.1. P-dataset consistency at the national scale

Considering all months, notable differences in P-dataset performance are observed at the national scale (Fig. 4). For example, KGE scores range between 0.34 (MERRA2-LND) and 0.86 (WFDEI-GPCC). This KGE score range can be partially explained by CC and Bias score variations (Eq. 4). **CC** scores range between 0.67 for WFDEI-CRU and 0.95 for MSWEP v.2.2. The lowest KGE scores are associated with the P-dataset that overestimates monthly precipitation estimates (MERRA2-FLX, ERA-5 LAND, IMERG-ER v.6, IMERG-FR v.6and MERRA2-LND). On the contrary, with a Bias value close to 0, the WFDEI-CRU KGE low value (0.53) is related to its low CC score value (0.67).

With a KGE (CC) score above 0.80 (0.90), the five best P-datasets (considering all months) are WFDEI-GPCC, CMORPH-BLD v.1, MSWEP v.2.2, CHIRPS v.2 and GSMaP-Adj v.6 (Fig. 4). Their elevated CC scores and minor variance demonstrate their ability to represent precipitation variations. Additionally, WFDEI-GPCC, CMORPH-BLD v.1, (and) MSWEP v.2.2 (< 45 mm month^{-1}) have the lowest RMSE scores, despite a positive median bias.

At the seasonal level and in terms of KGE and CC, all P-datasets that were studied are more reliable during the wet season than the dry season. During the dry season, only four P-datasets have a KGE > 0.50 (WFDEI-GPCC, CMORPH-BLD v.1, MSWEP v.2.2, and CHIRPS v.2), and some P-datasets even have a negative KGE score (ERA5, IMERG datasets, MERRA2-LND, PRISM, PERSIANN-CDR, TAMSAT v.3, and WFDEI-CRU). Globally, the error metrics tend to scale with the climatology: the Bias remains close to 0 for all P-datasets during the dry season, and Biais and RMSE are higher during the wet season since the precipitation amount is higher (Fig. 1).

Overall, the top five ranking P-datasets observed for all months in term of KGE (WFDEI_GPCC, CMORPH_BLD v.1, MSWEP v.2.2, GSMaP_Adj v.6, CHIRPS v.2) are similar during both wet (WFDEI_GPCC, CMORPH_BLD v.1, MSWEP v.2.2, TMPA_Adj v.7, CHIRPS v.2) and dry seasons (MSWEP v.2.2, WFDEI_GPCC, CHIRPS v.2, CMORPH_BLD v.1, GSMaP_Adj v.6).

4.1.2. P-dataset consistency at the climatic region level

We evaluated P-datasets at the regional scale for each climatic region. Results are detailed in Annex B.1. Overall, the P-datasets perform best in the dry region and worst in the humid region. In the dry region, KGE scores range from 0.52 (IMERG-LR v.6) to 0.93 (GSMaP- Adj v.6). More variability is observed in the other regions. In the semi-humid region, KGE scores range from -0.68 (MERRA2-FLX) to 0.93 (ARC-2). In the semi-arid region, KGE scores range from 0.12 (MERRA2-LND) to 0.90 (CHIRP v.2). In the humid region, KGE scores range from -0.21 (TAMSAT v.3) to 0.92 (WFDEI-GPCC). The most efficient P-dataset varies according to the climatic region under consideration: GSMaP-Adj v.6, ARC-2, CHIRP v.2, and WFDEI-GPCC are the most reliable P-datasets for dry, semi-humid, semi-arid, and humid regions, respectively.

P-dataset performance also varies within each climatic region (see KGE map on Fig. 5 and RMSE and CC maps in Appendix C.1 and Appendix C.2). For example, WFDEI-GPCC shows higher KGE scores in the northern than in the southern part of the semi-arid region (Fig. 5). Similarly, CMORPH-BLD v.1 (GSMaP-RT v.6) is more reliable in the northern part than in the southern part of the humid (dry) region. These results highlight the geographical variability of P-dataset performance. However, the low density of rain gauge stations prevents us from drawing hard conclusions at the local scale for each SPP.

The "best product" map (Fig. 5 lower right corner) shows the P-dataset with the highest KGE score at each rain gauge location. This map shows that product reliability varies across the island. Ideally, eight P-datasets should be considered to obtain the most reliable monthly precipitation estimates across Madagascar.

The following section illustrates the ensemble-merging approach using this complementarity to create a more reliable product for the entire island.

4.2. Ensemble-Merging approach

4.2.1. Calibrating a Merged Gridded Precipitation Dataset (MPD)

The weighted average process has been used to create a more a more highly performing P-dataset. For all climatic regions, consideration of the 21 P-datasets in the development of both MDP_{sa} and MDP_{wa} was in no case more satisfactory than the best P-dataset based on the KGE value (Fig. 6a, b, c, d). In fact, inclusion of P-datasets that are too different from observed values lowers the performance of both MPD_{aa} and MPD_{wa}. Consequently, an optimal number of P-datasets has been identified for each climatic region. Because MPD_{wa} systematically shows higher KGE than MPD_{aa}, this number was highlighted only for MPD_{wa}. The best combination includes two P-datasets in the humid region (CMORPH-BLD v.1 and WFDEI-GPCC, Fig. 6a) and in the dry region (CMORPH-BLD v.1

Table 2
scores of the five best products for (a) with Dataset-1 (24 rain gauge stations, period 2011–2016) and (b) with Dataset-2 (Five rain gauge stations, period 2000–2016). Bias is in mm month^{-1} and RMSE is in mm month^{-1}.

(a)	ALL YEAR				DRY SEASON				WET SEASON			
	KGE	CC	Bias	RMSE	KGE	CC	Bias	RMSE	KGE	CC	Bias	RMSE
MPD_{wa}	**0.93**	**0.98**	**2.3**	**22.8**	**0.72**	**0.89**	0.4	**7.7**	**0.90**	**0.97**	**1.3**	**32.3**
WFDEI GPCC	0.86	0.94	8.5	43.7	0.54	0.84	4.5	13.8	0.82	0.91	12.9	54.3
CMORPH BLD v.1	0.82	0.94	4.2	37.9	0.51	0.78	1.1	14.1	0.82	0.93	5.8	51.9
MSWEP v.2.2	0.81	0.95	2.5	39.6	0.66	0.84	**-0.2**	8.9	0.82	0.92	5.1	53.5
CHIRPS v.2	0.80	0.91	2.7	61.9	0.52	0.72	1.1	10.3	0.75	0.85	3.7	80.7

(b)	ALL YEAR				DRY SEASON				WET SEASON			
	KGE	**CC**	**Bias**	**RMSE**	**KGE**	**CC**	**Bias**	**RMSE**	**KGE**	**CC**	**Bias**	**RMSE**
MPD_{wa}	**0.90**	**0.95**	1.1	**36.8**	0.41	0.79	-4.8	**15.2**	**0.85**	**0.91**	**0.2**	**51.0**
CHIRP v.2	0.88	0.92	**0.7**	52.0	0.33	**0.82**	-3.0	22.9	0.81	0.83	4.3	71.4
ARC-2	0.87	0.90	-2.7	49.9	0.30	0.54	-7.3	23.8	0.78	0.81	5.4	68.0
CHIRPS v.2	0.83	0.91	-0.6	51.1	0.39	0.81	**-1.4**	17.9	0.77	0.80	1.5	71.9
CMORPH BLD v.1	0.81	0.92	3.3	47.9	**0.73**	0.79	1.7	18.5	0.75	0.80	11.0	66.9

and GSMaP-Adj v.6, Fig. 6c), five for the semi-humid region (ARC-2, MSWEP v.2.2, CMORPH-BLD v.1, WFDEI-GPCC, and GSMaP_Adj v.6, Fig. 6b), and six P-datasets for the semi-arid region (CHIRP v.2, CHIRPS v.2, TAMSAT v.3, MSWEP v.2.2, TMPA_RT v.7, and GSMaP_Adj v.6, Fig. 6d).

For each climate region, the use of a weighted average for MPD development results in a higher KGE score than using either the arithmetic average of P-datasets (Fig. 6a, b, c, d) or the most efficient P-dataset. For example, in the humid region, KGE values of 0.94, 0.90, and 0.92 are obtained for the MPD_{wa}, MPD_{aa} and for the best product (WFDEI-GPCC), respectively. MPDwa (solid line in Fig. 6e, f, g, h) is near rain gauge values (dotted line) for all seasons and in all regions. This indicates its ability to reproduce precipitation peaks in wet periods and the absence of precipitation in dry periods.

4.2.2. Validating MPD reliability

As expected, the MPDwa has a small Bias (2.3 mm month-1) and RMSE (22.8 mm month-1) and exhibits the highest KGE (0.93) and CC (0.98) scores for the calibration period (2011–2016, Dataset-1). However, to validate our methodology, this result must be tested against other data. Therefore, Dataset-2 (five stations for the period 2000–2016) was used to validate our results. For the validation period (2000–2016) and considering all months, MDP_{wa} is the best product in terms of KGE (0.89, see Table 2) and CC (0.95). The same result applies to wet season months (Table 2). During the dry season, the MDP_{wa} has a favorable CC score (0.79) but tends to underestimate precipitation and show a lower KGE score (0.41).

5. Discussion

5.1. Benefit of the GPM-based precipitation estimate over TRMM-based precipitation estimates

The transition from TRMM-based P-datasets to GPM-based datasets was not successful across Madagascar. IMERG P-datasets were developed as a successor to TMPA P-datasets, which are no longer produced (Zhu et al., 2020). Despite improvements in satellite data and retrieval algorithms between the two datasets, the results show no significant improvement. TMPA-Adj v.7 (TMPA-RT v.7) monthly precipitation estimates are more reliable than that derived from its counterpart IMERG-FR v.6 (IMERG-ER and LR v.6). This result is consistent with a study comparing P-dataset reliability across the Jurua basin of the Amazon region (Satgé et al., 2021). The authors found that TMPA-Adj v.7 (TMPA-RT v.7) daily precipitation estimates were more reliable than that derived from IMERG-FR v.6 (IMERG-ER and LR v.6). However, another study conducted over Nigeria, shows the opposite result, with IMERG-FR v.6 (IMERG-ER and LR v.6) monthly precipitation estimates that are more accurate than TMPA-Adj v.7 (TMPA-RT v.6) (Nwachukwu et al., 2020). Similarly, CMORPH P-datasets also outperformed IMERG P-datasets for both gauge-adjusted and non-adjusted versions, which is consistent with previous observations made across the Jurua Amazon basin (Nigeria) (Satgé et al., 2021, Nwachukwu et al., 2020). The contradictory results regarding the performance of the oldest and the newest P-datasets obtained across different regions confirms the need for a P-dataset reliability assessment prior to their use over a specific region.

Finally, versions that include a gauge-based adjustment outperformed non-gauge-adjusted versions. For example, CMORPH-BLD v.1, GSMaP-Adj v.6, TMPA-Adj v.7 and IMERG-FR v.6 show higher KGE scores than CMORPH-CRT v.1, GSMaP-RT v.6, TMPA-RT v.7 and IMERG-ER and-LR v.6, respectively (Fig. 2). This result shows the benefit of such adjustment and the need to maintain traditional gauge observations despite their low spatial coverage and consequent difficulty in adequately capturing spatial precipitation variability.

Table A1
Meteorological stations characteristics for the period 2011–2016.

Name	Long.	Lat.	Alt. (m)	Temporal coverage	Climatic region	GSOM	GSOD	GPCC
Ampanihy	44.75	-24.65	248	1981–2018	Semi-arid	-	-	No
Ankazoabo Sud	44.55	-22.45	882	1981–2018	Semi-arid	-	-	No
Antananarivo	47.53	-18.90	1260	2011–2018	Semi-humid	1977–2023	2011–2019	Yes
Antsirabe	47.07	-19.87	1542	2011–2018	Semi-humid	1977–2022	1974–2015	Yes
Antsiranana	49.30	-12.35	115	2011–2018	Dry	1977–2023	-	Yes
Antsohihy	47.96	-14.88	9	2011–2018	Dry	-	1987–2020	Yes
Besalampy	44.48	-16.75	46	2011–2018	Dry	-	1973–2007	Yes
Betioky Sud	44.35	-23.75	240	1981–2018	Semi-arid	-	-	No
Ejeda	44.55	-24.35	249	1981–2018	Semi-arid	-	-	No
Farafangana	47.83	-22.80	3	2011–2018	Humid	1977–2005	1973–2015	Yes
Fianarantsoa	47.10	-21.45	1136	2011–2018	Semi-humid	1977–2022	1973–2020	No
Ivato	47.48	-18.80	1268	2011–2018	Semi-humid	-	1973–2020	Yes
Mahajanga	46.35	-15.66	17	2011–2018	Dry	-	-	Yes
Maintirano	44.03	-18.05	26	2011–2018	Dry	1978–2015	1973–2017	Yes
Mananjary	48.36	-21.20	3	2011–2018	Humid	1977–2015	1973–2015	Yes
Morombe	43.36	-21.75	2	2011–2018	Semi-arid	1980–2004	1973–2004	Yes
Morondava	44.31	-20.28	4	2011–2018	Semi-arid	1977–2018	1973–2020	Yes
Nosy Be	48.31	-13.32	17	2011–2018	Semi-humid	-	-	Yes
Ranohira	45.40	-22.55	763	2011–2018	Semi-arid	2006–2020	1973–2020	Yes
Sainte-Marie	49.81	-17.08	4	2011–2018	Humid	1987–2023	1973–2020	Yes
Sakaraha	44.55	-24.95	217	1981–2018	Semi-arid	-	-	No
Sambava	50.16	-14.28	68	2011–2018	Humid	1977–2022	1973–2020	Yes
Toamasina	49.40	-18.11	9	2011–2018	Humid	1977–2022	1973–2020	Yes
Toliary	43.73	-23.38	9	1980–2018	Semi-arid	1977–2022	1973–2019	No

Table A2
Meteorological stations characteristics for the period 2000–2016.

Name	Long.	Lat.	Alt. (m)	Temporal coverage	Climatic region	GSOM	GSOD	GPCC
Andranoma-nelatra	47.11	-19.78	1621	2002–2020	Semi-humid	-	-	No
Befandefa	43.25	-22.15	9	1981–2018	Semi-arid	-	-	No
Ivory	46.42	-19.55	926	2005–2020	Semi-humid	-	-	No
Mandrare	46.35	-25.05	13	1981–2018	Semi-arid	-	-	No
Manombo Sud	43.45	-22.95	10	1981–2018	Semi-arid	-	-	No

GSOD = Global Summary Of the Day (available at https://www.ncei.noaa.gov/maps/daily/?layers=0001, accessed on 11/03/2023),
GSOM = Global Summary of the Month (available at https://www.ncei.noaa.gov/maps/monthly/, accessed on 11/03/2023)
GPCC = Global Precipitation Climatology Center (available at: https://kunden.dwd.de/GPCC/Visualizer_intro?p_dataset=monitor10_v2022&p_coast=1&p_art=3&p_out=0&p_season=1&p_scale=1.4&p_year=2020&p_as=0&p_zoom=1&p_area=0&p_lonmin=--180.&p_lonmax=%2B180.&p_latmin=−90.&p_latmax=%2B90.&p_color=1&p_proj=1, accessed on 11/03/2023)

5.2. Benefit of multiple data combinations for precipitation estimates?

Most of the best P-datasets combine multiple data sources. For example, MSWEP v.2.2 and CHIRPS v.2, which combine satellite, reanalysis, and gauge information, are among the 5 most reliable P-datasets. MSWEP v.2.2 was built to optimize existing precipitation data sources by combining them, in particular to mitigate impacts of orographic constraints and low rain gauge density (Beck et al., 2017). MSWEPv.2.2 merges CMORPH, GSMAP, TMPA, and ERA-Interim data in addition to other data sources. The results validate the multi-source approach as shown by the fact that MSWEP v.2.2 is among the best performing products across all regions of Madagascar (Appendix B1). Similarly, CHIRPS v.2, which combines information from TMPA, MERRA2, ERA5, and gauge information, ranked in the 5 most reliable P-datasets (Fig. 2).

Because soil moisture is a result of precipitation events, efforts have been made to retrieve precipitation data from soil moisture estimates in a bottom-up approach (Brocca et al., 2013) instead of the approach using cloud properties. These approaches have proven efficient in flat semi-arid to arid regions where satellite soil moisture data are more reliable due to low interference from topography and/or vegetation cover (Satgé et al., 2021). In this context, the PRISM (Precipitation Inferred from Soil Moisture) product seeks to improve IMERG-ER estimates by using soil moisture data from the Soil Moisture and Ocean Salinity (SMOS) satellite (Pellarin et al., 2022). The results show an obvious improvement in KGE with the PRISM product, especially in dry and semi-arid regions. These results are consistent with the work of Pellarin et al. (2022) in West Africa. PRISM performs better in dry regions because the soil moisture constraint prevents an overestimation of the number of rainfall events. Even so, the low reliability of SMOS in coastal regions limits the performance of PRISM over an island.

Combining satellite-based data and gauge-based data enhance P-dataset performance. As an example, there is a clear enhancement in precipitation estimates between CHIRP v.2 (no gauge information) and CHIRPS v.2 at the national scale and especially in the wet region (KGE_{CHIRP} = 0.5, KGE_{CHIRPS} = 0.65). Similarly, WFDEI-GPCC performs better than WFDEI-CRU. This difference can be

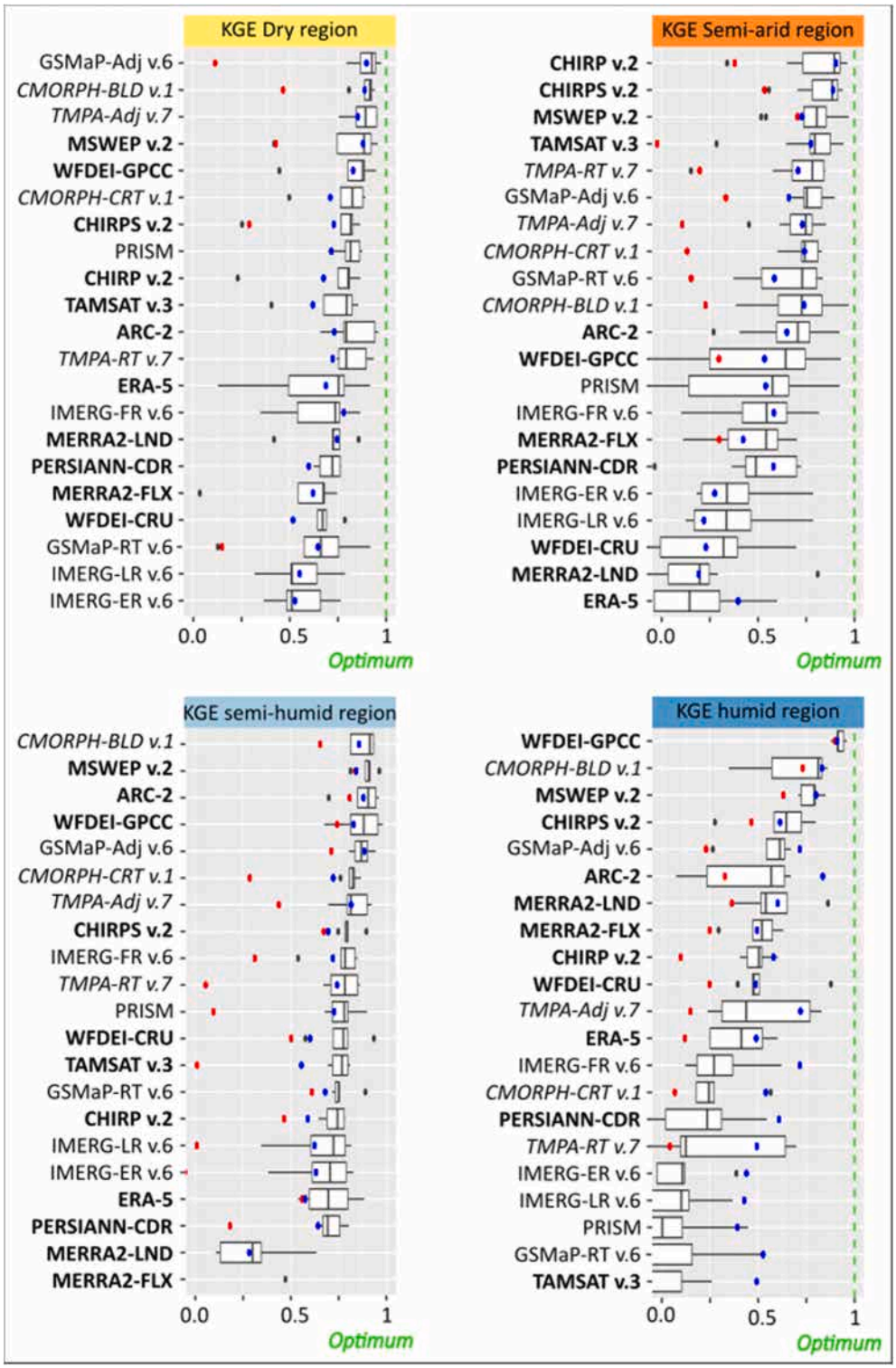


Fig. B1. Boxplot of KGE scores at the regional level.

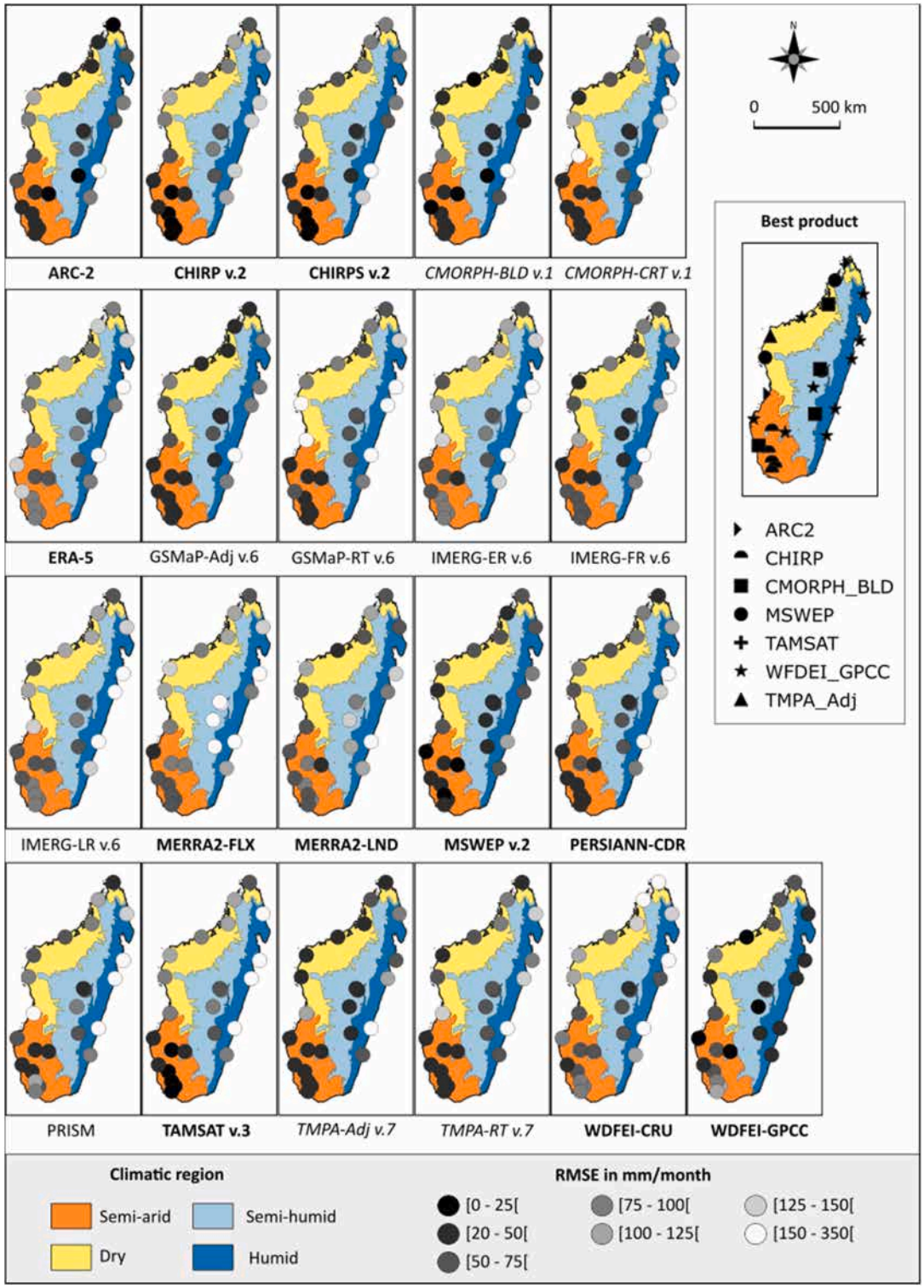


Fig. C1. Map of the regional RMSE scores for the period 2011–2016.

explained by the number of rain gauge stations used: four stations in the CRU dataset (Harris et al., 2020), about 20 in the GPCC dataset (Schneider et al., 2022) for Madagascar. This means that WFDEI products are highly dependent on rain gauge network density, as previously observed by Chen et al. (2020).

To benefit from the importance of gauge information, ARC2 and TAMSAT v.3 have been specifically developed to use as much

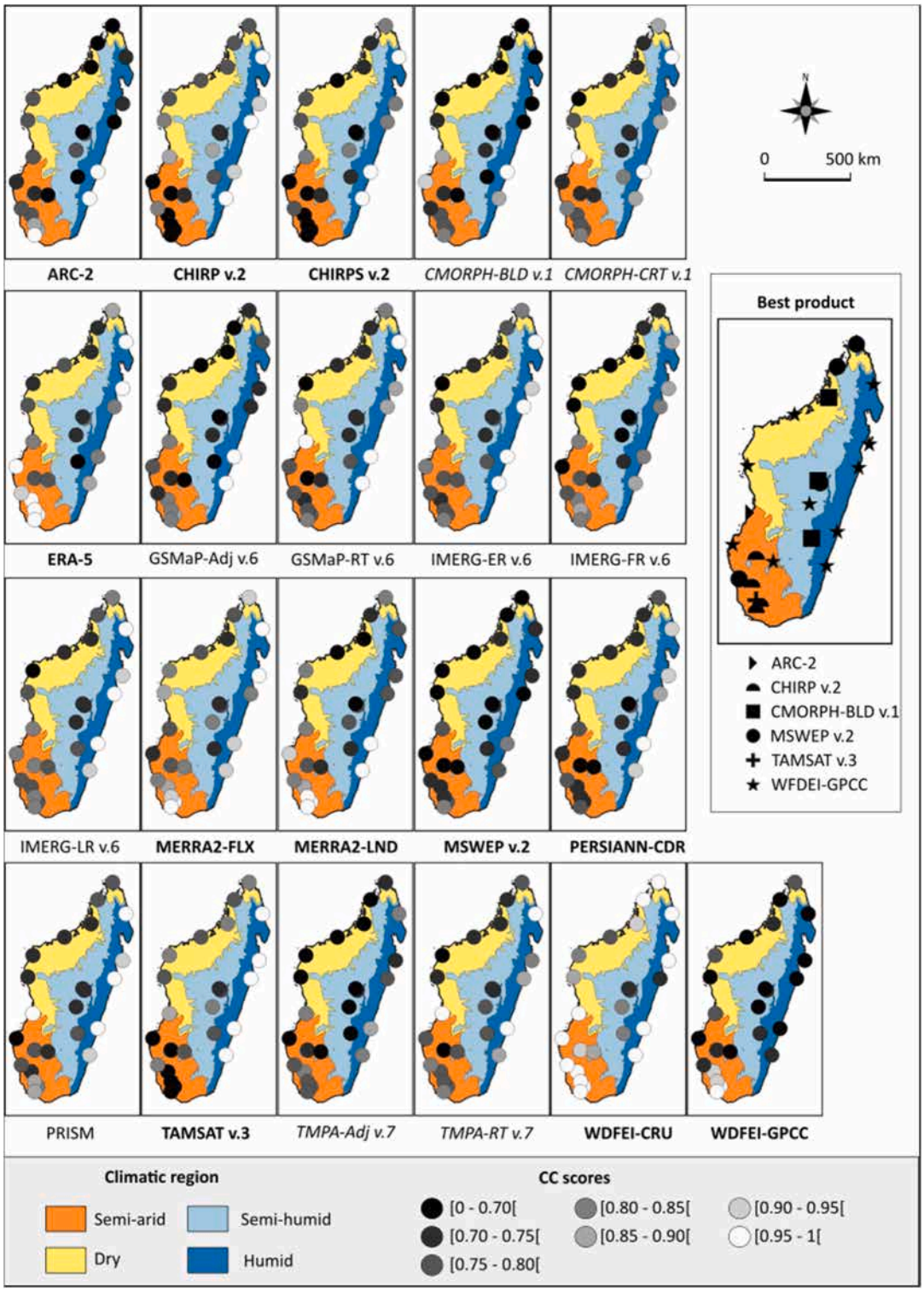


Fig. C2. Map of the regional CC scores for the period 2011–2016.

gauge information as possible to estimate precipitation across Africa. ARC2 was specifically developed to improve rainfall monitoring needed/necessary for water management, agriculture, and food security on the African continent. Novella and Thiaw (2013) showed that this product estimates droughts fairly accurately but underestimates rainfall during the dry season (June-October) and in coastal and mountainous regions. Our study results confirm this bias, with a correlation coefficient close to 0.8 in the dry season, with a

negative bias, and a low KGE score in the wet region. TAMSAT v.3, designed to estimate low precipitation from a drought monitoring perspective, has been shown to underestimate precipitation (Maidment et al., 2014, 2017). This underestimation results in better scores in semi-arid regions than in wet regions (negative KGE), and an overall negative bias.

5.3. Ensemble-merging approach advantages and limitations

Multiple methods have been developed to explore the merging of P-datasets, but none is widely implemented, and none has been evaluated in Madagascar. Consequently, a weighted averaging approach was proposed in this paper, based on the principles of already existing approaches. The resulting MPD_{wa} product performed better than other products for the calibration and validation periods. The uniqueness of this method is the selection of P-datasets based on their observed performance over each climatic region. This approach differs from previous studies that include only a limited number of P-datasets in the selection and merging process (Mastrantonas et al., 2019; Rahman et al., 2018; Zhang et al., 2021). This broad selection was motivated by the variability in P-dataset performance specific to the climatic regions of Madagascar (Fig. 5). Other spatial characteristics could have been used to constrain the MPD weightings (e.g. altitude, slope). However, our tests suggested that slope and altitude alone are not sufficient to optimize the weightings, perhaps due to the effects of trade winds. Therefore, we have chosen to use climatic regions as the spatial variation factor, which indirectly includes altitude and slope.

The results show a high variability according to the season. This observation implies that it may be advisable to optimize MPD weightings according to the season. This weighting method would require expertise in seasonal variations, which could be the focus of a new study. The weightings could be computed using a Weighted Average Least Squares (WALS) to combine P-datasets with weightings adjusted according to the seasons, as previously tested by Rahman et al. (2020) across Pakistan.

Using MPD makes it possible to estimate uncertainty bounds. This uncertainty is related to the dispersion of the values of the various P-datasets that were used to establish MPD (dark gray area in Fig. 6). The uncertainty could be estimate based on the value range or on the standard variation of the products selected for the MPD values. Knowledge of the uncertainty is essential in being able to propagate it throughout the study. For example, if the precipitation and evapotranspiration uncertainties are known, it will be possible to estimate the uncertainty on the P-ET term, which is a proxy for aquifer recharge in these semi-arid areas without rivers.

6. Conclusion

The proliferation of P-datasets makes precipitation data available even in regions with low-density rain gauge networks, such as Madagascar. However, P-dataset performance varies in time and space. It is therefore essential to choose a P-dataset that is robust enough to cope with the various climatic contexts and seasons observed in Madagascar. In this context, this study assesses for the first time the reliability of 21 P-datasets at the monthly time step across Madagascar. The results are summarized below:

- At the national scale, WDFEI-GPCC, CMORPH-BLD v.1, and MSWEP v.2.2 are the most reliable P-datasets. With a spatial resolution of 0.1° and temporal coverage of 44 years (1979-present), MSWEP v.2.2 offers encouraging prospects for precipitation trend analysis across Madagascar provided that this product meets other criteria specific to trend analysis such as continuity of input data.
- P-dataset reliability varies in space according to the climatic region under consideration. GSMaP-Adj v.6, CHIRP v.2, CMORPH-BLD v.1, and WFDEI-GPCC are the most reliable P-datasets for dry, semi-arid, semi-humid, and humid regions, respectively.
- The ensemble-merging approach proposed in this study takes advantage of the variability of P-dataset reliability in space (i.e., climatic region). The MPD_{wa} P-dataset developed according to this method provides precipitation estimates that are more reliable than all other P-datasets that were considered. This was true for all climatic regions.
- The two-step method (P-dataset reliability assessment and ensemble-merging approach) presented in this paper could be tested at other time steps or in other regions that have available rain gauge measurements. This method makes it possible to take advantage of all available P-datasets to obtain more reliable precipitation estimates across the study area.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Simon Carriere reports financial support was provided by French Space Agency.

Acknowledgements

This project involved the use of precipitation data acquired and shared by: DGM (Direction générale de la météorologie) of Madagascar and CIRAD (Centre de coopération internationale en recherche agronomique pour le développement) of France. Financial support for part of the study was provided by CNES APR (SuFECiS project) and by the SIWA Foundation. This project used satellite-based precipitation datasets made available by the following organizations:

- ARC2: https://www.icpac.net/data-center/arc2/
- CHIRP: https://data.chc.ucsb.edu/products/CHIRP/

- CHIRPS: Funk, C.C., Peterson, P.J., Landsfeld, M.F., Pedreros, D.H., Verdin, J.P., Rowland, J.D., Romero, B.E., Husak, G.J., Michaelsen, J.C., and Verdin, A.P., 2014, A quasi-global precipitation time series for drought monitoring: U.S. Geological Survey Data Series 832, 4 p. ftp://chg-ftpout.geog.ucsb.edu/pub/org/chg/products/CHIRPS-2.0/docs/USGS-DS832. CHIRPS.pdf, data access: https://data.chc.ucsb.edu/products/CHIRPS-2.0/
- CMORPH-BLD & CMORPH-CRT: https://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/, Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydromet., 5, 487-503.
- ERA-5 Land: Copernicus Climate Change Service, Climate Data Store, (2021): ERA5-Land monthly averaged data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 15-09-2021), 10.24381/cds.68d2bb30
- GSMAP-RT & GSMAP-ADJ: ' Global Rainfall Map (GSMaP) by JAXA Global Rainfall Watch' was produced and distributed by the Earth Observation Research Center, Japan Aerospace Exploration Agency. See Kubota, T., K. Aonashi, T. Ushio, S. Shige, Y. N. Takayabu, M. Kachi, Y. Arai, T. Tashima, T. Masaki, N. Kawamoto, T. Mega, M. K. Yamamoto, A. Hamada, M. Yamaji, G. Liu and R. Oki 2020: Global Satellite Mapping of Precipitation (GSMaP) products in the GPM era, Satellite precipitation measurement, Springer, https://doi.org/10.1007/978-3-030-24568-9_20. Available at https://sharaku.eorc.jaxa.jp/GSMaP_NOW/index.htm Accessed on 15-09-2021
- IMERG-ER: Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, Jackson Tan (2019), GPM IMERG Early Precipitation L3 Half Hourly 0.1° x 0.1° V06, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15-09-2021, 10.5067/GPM/IMERG/3B-HH-E/06. Available at: https://disc.gsfc.nasa.gov/
- IMERG-FR: Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, Jackson Tan (2019), GPM IMERG Final Precipitation L3 1 month 0.1° x 0.1° V06, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15-09-2021 10.5067/GPM/IMERG/3B-MONTH/06. Available at https://disc.gsfc.nasa.gov/
- IMERG-LR: Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, Jackson Tan (2019), GPM IMERG Late Precipitation L3 1 day 0.1° x 0.1° V06, Edited by Andrey Savtchenko, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15-09-2021, 10.5067/GPM/IMERGDL/DAY/06. https://disc.gsfc.nasa.gov/
- MERRA-2-FLX: Global Modeling and Assimilation Office (GMAO) (2015), MERRA-2 tavgM_2d_flx_Nx: 2d, Monthly mean, Time-Averaged, Single-Level, Assimilation, Surface Flux Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15/09/2021. 10.5067/0JRLVL8YV2Y4
- MERRA-2-LND: Global Modeling and Assimilation Office (GMAO) (2015), MERRA-2 tavgM_2d_lnd_Nx: 2d,Monthly mean, Time-Averaged, Single-Level, Assimilation, Land Surface Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15/09/2021. 10.5067/8S35XF81C28F
- MSWEP: Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. M., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F. MSWEP V2 global 3–hourly 0.1° precipitation: methodology and quantitative assessment Bulletin of the American Meteorological Society 100(3), 473–500, 2019. Available at http://www.gloh2o.org/mswep/, accessed on 15/09/2021
- PRISM: Pellarin Thierry, Kerr Yann, Roman-Cascon Carlos, & Zoppis Alexandre. (2020). PrISM satellite rainfall product (2010-2020) based on SMOS soil moisture measurements in Africa (3 h, 0.25°) [PrISM_Rainfall_SMOS-L3SM_2011-2021_010d]. Zenodo. https://doi.org/10.5281/zenodo.4769825 Available at https://zenodo.org/record/4769825, accessed on 10/09/2022.
- PERSIANN-CDR: - Ashouri, H., K.L. Hsu, S. Sorooshian, D.K. Braithwaite, K.R. Knapp, L.D. Cecil, B.R. Nelson, and O.P. Prat, 2015:
- PERSIANN-CDR: 1) Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. Bull. Amer. Meteor. Soc., 96, 6983 doi: https://doi.org/10.1175/BAMS-D-13-00068.1 2) Nguyen, P., E.J. Shearer, H. Tran, M. Ombadi, N. Hayatbini, T. Palacios, P. Huynh, G. Updegraff, K. Hsu, B. Kuligowski, W.S. Logan, and S. Sorooshian, The CHRS Data Portal, an easily accessible public repository for PERSIANN global satellite precipitation data, Nature Scientific Data, Vol. 6, Article 180296, 2019. doi: https://doi.org/10.1038/sdata.2018.296. Available at https://chrsdata.eng.uci.edu/ (accessed on 15/09/2021).
- TAMSAT: 1) Maidment, R. I., D. Grimes, E. Black, E. Tarnavsky, M. Young, H. Greatrex, R. P. Allan et al. (2017). A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa Nature Scientific Data 4: 170063 DOI:10.1038/sdata.2017.63. 2) Tarnavsky, E., D. Grimes, R. Maidment, E. Black, R. Allan, M. Stringer, R. Chadwick, F. Kayitakire (2014). Extension of the TAMSAT Satellite-based Rainfall Monitoring over Africa and from 1983 to present Journal of Applied Meteorology and Climate DOI 10.1175/JAMC-D-14-0016.1 3) Maidment, R., D. Grimes, R.P.Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling and E. Black (2014). The 30 year TAMSAT African Rainfall Climatology And Time series (TARCAT) data set Journal of Geophysical Research DOI: 10.1002/2014JD02192. Available at http://www.tamsat.org.uk/data (accessed on 15/09/2021).
- TMPA-ADJ: Huffman, G.J., D.T. Bolvin, E.J. Nelkin, and R.F. Adler (2016), TRMM (TMPA) Precipitation L3 1 day 0.25° x 0.25° V7, Edited by Andrey Savtchenko, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15/09/2021, 10.5067/TRMM/TMPA/DAY/7. Available at https://disc.gsfc.nasa.gov/.
- TMPA-RT: Goddard Earth Sciences Data and Information Services Center (2016), TRMM (TMPA-RT) Near Real-Time Precipitation L3 1 day 0.25° x 0.25° V7, Edited by Andrey Savtchenko, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 15/09/2021. 10.5067/TRMM/TMPA/DAY-E/7. Available at https://disc.gsfc.nasa.gov/.
- WFDEI-GPCC & WFDEI-CRU: Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo. 2018. *The WFDEI Meteorological Forcing Data*. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/486 N-8109 Accessed 15/09/2021. Available at https://rda.ucar.edu/datasets/ds314.2/dataaccess/.

The authors wish to acknowledge the peer reviewer for the constructive and helpful comments and suggestions.

Appendix

See Tables A1 and A2.

Reliability assessment for 20 P-datasets at the regional level, using monthly data from 24 meteorological stations (period 2011–2016). The dotted green line is the optimum for each indicator. The right and left edges of the boxes represent the 25th and 75th percentile values, respectively. The blue dot is the median value for the wet season, and the red dot is the median value for the dry season. The P-datasets are ranked by the KGE median value Fig. B1.

At the regional level, P-datasets performances vary according to the season. In the semi-arid region, the best P-datasets are mostly the same for all seasons, except for TAMSAT v.7, which has a high global KGE, but a negative one for the dry season. In this region, CHIRP v.2 and CHIRPS v.2 present high KGE scores, especially for the wet season. In the semi-humid region, the best P-dataset being the same for the dry and the wet seasons. Among the five best P-datasets, only CMORPH-BLD v.1 shows significantly lower results during the dry season. In the humid region, WFDEI-GPCC and CMORPH-BLD v.1 present high scores for all seasons. MSWEP v.2 is also efficient but with a lower score for the dry season. Several P-datasets fail to reproduce precipitation amounts for the dry season, but present high KGE scores for the wet season (ARC-2, IMERG-FR v.6, GSMaP- Adj v.6). In the dry region, MSWEP v.2 and CMORPH-BLD v.1 are among the best for all seasons. GSMaP- Adj v.6 and TMPA- Adj v.7 have high scores for the wet season and low scores for the dry season Fig. C1.

RMSE scores according to the climatic regions. RMSE scores obtained with monthly data for the period 2011–2016 are shown per station for the 20 P-datasets. The best product map indicates the product with the highest KGE for each rain gauge station.

The RMSE scores are mostly better in the semi-arid region. Only WFDEI-GPCC achieves a small RMSE in the humid region. CMORPH-BLD v.1 has good RMSE scores for most of the island except the west coast. MSWEP v.2 has balanced scores Fig. C2.

CC scores according to the climatic regions. RMSE scores obtained with monthly data for the period 2011–2016 are shown per station for the 20 P-datasets. The best product map indicates the product with the highest KGE for each rain gauge station.

The CC scores at the regional level are in line with the KGE scores.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2023.101400.

References

Beck, H.E., Van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., De Roo, A., 2017. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21 (1), 589–615. https://doi.org/10.5194/hess-21-589-2017.

Carrière, S.D., Health, T., Rakotomandrindra, P.F.M., Ollivier, C., Rajaomahefasoa, R.E., Rakoto, H.A., Lapègue, J., Rakotoarison, Y.E., Mangin, M., Kempf, J., Razakamanana, T., Chalikakis, K., 2021. Long-term groundwater resource observatory for Southwestern Madagascar. Hydrol. Process. 35 (3), 1–6. https://doi.org/10.1002/hyp.14108.

Chen, S., Liu, B., Tan, X., Wu, Y., 2020. Inter-comparison of spatiotemporal features of precipitation extremes within six daily precipitation products. Clim. Dyn. 54 (1–2), 1057–1076. https://doi.org/10.1007/s00382-019-05045-z.

Collier, C.G., 1986. Accuracy of rainfall estimates by radar, part I: calibration by telemetering raingauges. J. Hydrol. 83 (3–4), 207–223. https://doi.org/10.1016/0022-1694(86)90152-6.

Cornet, A. (1974). Essai cartographique bioclimatique à Madagascar, carte à 1/2′000′000 et notice explicative.

Depraetere, C., Gosset, M., Ploix, S., Laurent, H., 2009. The organization and kinematics of tropical rainfall systems ground tracked at mesoscale with gages: First results from the campaigns 1999-2006 on the Upper Ouémé Valley (Benin). J. Hydrol. 375 (1–2), 143–160. https://doi.org/10.1016/j.jhydrol.2009.01.011.

FEWSNET. (2022). Famine Early Warning System Network - Madagascar. WWW Document. https://fews.net/fr/southern-africa/madagascar.

Gehne, M., Hamill, T.M., Kiladis, G.N., Trenberth, K.E., 2016. Comparison of global precipitation estimates across a range of temporal and spatial scales. J. Clim. 29 (21), 7773–7795. https://doi.org/10.1175/JCLI-D-15-0618.1.

Grimes, D.I.F., 2008. An ensemble approach to uncertainty estimation for satellite-based rainfall estimates. Hydrol. Model. Water Cycle 145–162. https://doi.org/10.1007/978-3-540-77843-1_7.

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

Harris, I., Osborn, T.J., Jones, P., Al, E., 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.

Huffman, G.J. (2020). 2 October 2020 TMPA-to-IMERG Transition 1. 2019(October), 1–5.

Kidd, C., Huffman, G., 2011. Global precipitation measurement. Meteorol. Appl. 18 (3), 334–353. https://doi.org/10.1002/met.284.

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.

Lebel, T., Cappelaere, B., Galle, S., Hanan, N., Kergoat, L., Levis, S., Vieux, B., Descroix, L., Gosset, M., Mougin, E., Peugeot, C., Seguis, L., 2009. AMMA-CATCH studies in the Sahelian region of West-Africa: an overview. J. Hydrol. 375 (1–2), 3–13. https://doi.org/10.1016/j.jhydrol.2009.03.020.

Maidment, R.I., Grimes, D., Allan, R.P., ElenaTarnavsky, Stringer, M., Hewison, T., Roebeling, R., Black, E., 2014. The 30 year TAMSAT African rainfall climatology and time series (TARCAT) data set. J. Geophys. Res. Atmospheres *119*. https://doi.org/10.1002/2014JD021927.

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. https://doi.org/10.1038/sdata.2017.63.

Mastrantonas, N., Bhattacharya, B., Shibuo, Y., Rasmy, M., Espinoza-Dávalos, G., Solomatine, D., 2019. Evaluating the benefits of merging near-real-time satellite precipitation products: a case study in the Kinu basin region, Japan. J. Hydrometeorol. 20 (6), 1213–1233. https://doi.org/10.1175/JHM-D-18-0190.1.

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

Nwachukwu, P.N., Satge, F., Yacoubi, S., El, Pinel, S., Bonnet, M.P., 2020. From trmm to GPM: How reliable are satellite-based precipitation data across Nigeria. Remote Sens. 12 (23), 1–22. https://doi.org/10.3390/rs12233964.

Pellarin, T., Zoppis, A., Román-Cascón, C., Kerr, Y.H., Rodriguez-Fernandez, N., Panthou, G., Philippon, N., Cohard, J.M., 2022. From SMOS Soil Moisture to 3-hour Precipitation Estimates at 0.1∘ Resolution in Africa. Remote Sens. 14 (3) https://doi.org/10.3390/rs14030746.

Poccard-leclercq, I. (2000). Etude diagnostique de nouvelles données climatiques: les réanalyses. Exemples d′application aux précipitations en Afrique tropicale.

Rahman, K.U., Shang, S., Shahid, M., Wen, Y., 2020. An appraisal of dynamic bayesian model averaging-based merged multi-satellite precipitation datasets over complex topography and the diverse climate of Pakistan. Remote Sens. 12 (1) https://doi.org/10.3390/RS12010010.

Rahman, K.U., Shang, S., Zohaib, M., 2021. Assessment of merged satellite precipitation datasets in monitoring meteorological drought over pakistan. Remote Sens. 13 (9), 1–37. https://doi.org/10.3390/rs13091662.

Ramahaimandimby, Z., Randriamaherisoa, A., Jonard, F., Vanclooster, M., Bielders, C.L., 2022. Reliability of gridded precipitation products for water management studies: the case of the Ankavia River Basin in Madagascar. Remote Sens. 14 (16) https://doi.org/10.3390/rs14163940.

Randriatsara, H.H.H., Hu, Z., Xu, X., Ayugi, B., Sian, K.T.C.L.K., Mumo, R., Ongoma, V., 2022. Evaluation of gridded precipitation datasets over Madagascar. Int. J. Climatol., Dec. 2021, 1–19. https://doi.org/10.1002/joc.7628.

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, 1–47. https://doi.org/10.1016/j.jhydrol.2019.124412.

Schneider, U., Finger, P., Rustemeier, E., Ziese, M., & Hänsel, S. (2022). Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre (GPCC), Deutscher Wetterdienst, Offenbach a. M.

Sinclair, S., Pegram, G., 2005. Combining radar and rain gauge rainfall estimates using conditional merging. Atmos. Sci. Lett. 6 (1), 19–22. https://doi.org/10.1002/asl.85.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.L., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107. https://doi.org/10.1002/2017RG000574.

UNICEF. (2018). Bulletin d′information du cluster nutrition 2018-T1 - Résultats de la surveillance nutritionnelle dans huit districts du Sud de Madagascar. Antananarivo.

UNICEF. (2019). Bulletin d′information du cluster nutrition 2019-T1 - Résultats de la surveillance nutritionnelle dans huit districts du Sud de Madagascar. Antananarivo.

Ur Rahman, K., Shang, S., Shahid, M., Li, J., 2018. Developing an ensemble precipitation algorithm from satellite products and its topographical and seasonal evaluations over Pakistan. Remote Sens. 10 (11) https://doi.org/10.3390/rs10111835.

World Bank. (2020). Poverty headcount ratio at national poverty lines (% of population). WWW Document.

World Health Organization. (2015). Climate and health country profile 2015: Madagascar.

Zhang, L., Li, X., Zheng, D., Zhang, K., Ma, Q., Zhao, Y., Ge, Y., 2021. Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach. J. Hydrol. 594 (August 2020), 125969 https://doi.org/10.1016/j.jhydrol.2021.125969.

Zhu, S., Shen, Y., & Ma, Z. (2020). A New Perspective for Charactering the Spatio-temporal Patterns of the Error in GPM IMERG Over Mainland China. https://doi.org/https://doi.org/10.1029/2020EA001232.