



Evaluation of gridded precipitation products in the selected sub-basins of Lower Mekong River Basin

Santosh Dhungana¹ · Sangam Shrestha^{1,2} · Tuan Pham Van¹ · Saurav KC¹ · Ashim Das Gupta¹ · Thi Phuoc Lai Nguyen³

Received: 30 January 2022 / Accepted: 27 October 2022 / Published online: 15 November 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Hydrological and meteorological studies demand accurate, continuous, long-term, reliable, and uniformly distributed precipitation data. Considering low density rain gauges with incomplete data in developing nations, a plethora of gridded precipitation products (GPPs) have made their place as an alternative to rain gauge records. However, GPPs house inherent errors depending on the type of data, gauge density, gridding algorithm, etc. Hence, it is crucial to evaluate them prior to their application. This study evaluated monthly products of eight GPPs over 17 years (1998–2014) – Asian Precipitation Highly Resolved Observational Data Integration towards Evaluation data (APHRODITE), Climate Prediction Center (CPC), Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), Southeast Asian Observed dataset (SA-OBS), Climate Prediction Center Morphing Technique (CMORPH), The Tropical Rainfall Measuring Mission (TRMM)-daily products, Climate Research Unit (CRU), and Global Precipitation Climatology Center (GPCC). An entropy-based weight calculation for each statistical index and compromise programming was employed to rank the GPPs in the selected sub-basins (Nam Ngum River Basin, NRB, and Vietnam Mekong Delta, VMD) of the Lower Mekong Region (LMR) for mean and six extreme precipitation indices. The correlation coefficient (r), root mean square error (RMSE), skilled score (SS), and bias were the continuous statistical indices and probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI) were the categorical indices used in this study. In terms of capturing mean monthly precipitation, GPCC outweighed all other products for both the studied basins. However, APHRODITE ranked first for daily precipitation products based on compromise programming algorithm for NRB. APHRODITE consistently recorded r between 0.85 and 0.95, RMSE between 50 and 100 mm/month, and SS between 0.72 and 0.90 for the 5 observed stations. Similarly, in case of VMD, TRMM ranked first for the daily precipitation products with r between 0.8 and 0.95, RMSE between 50 and 70 mm/month, and SS between 0.56 and 0.9 when evaluated with 11 observed stations. The APHRODITE for NRB and TRMM for VMD can be used as alternate to gauge data for hydrological and meteorological studies.

Keywords Gridded precipitation products · Entropy method · Compromise programming · Evaluation

✉ Santosh Dhungana
santosh.dhung@gmail.com

¹ Water Engineering and Management (WEM), Civil and Infrastructure Engineering (CIE), School of Engineering and Technology (SET), Asian Institute of Technology (AIT), 12120 Pathum Thani, Thailand

² Stockholm Environment Institute (SEI), Asia Centre, Chulalongkorn Soi 64, Phyathai Road, 1033 Pathumwan, Bangkok, Thailand

³ Development Planning Management and Innovation (DPMI), Department of Development and Sustainability, School of Environment, Resources and Development (SERD), Asian Institute of Technology (AIT), 12120 Pathum Thani, Thailand

1 Introduction

Precipitation, a key environmental variable, finds itself in a multitude of applications making it central to scientific research such as hydroclimatic modelling, climate change, agriculture, and water resource modelling and management (Sun et al. 2018; Try et al. 2020). It is also noteworthy to mention that obtaining a high-quality and fine spatiotemporal resolution precipitation data is still a challenge. Direct gauge precipitation is considered the most reliable and accurate data source that reflects the precipitation at the Earth's surface. However, in many developing nations, availability of such gauge-based data is dire given their political-economic instability. Even though gauge data are available

for some, they may not be complete and are often costlier. Also, they house inhomogeneities attributed to instrumentation relocation, observation gaps and errors, data transmission, inability to represent single large event, etc. Weather radar could be an alternative to the gauge-precipitation, but it too demands wider coverage and calibration according to precipitation type. Further, it also could be the costlier and have limitations in capturing the precipitation in mountainous areas (Arshad et al. 2021; Lee 2006). To overcome these gaps, gridded precipitation products (GPPs) developed at semi-global to global scale with varying spatio-temporal resolution using satellite estimates, gauge-based interpolated products, and/or a blend of both have garnered wider attention from the scientific communities.

GPPs are developed drawing on a vast wealth of information from several sources and agencies (both government and non-government) further intensified by leveraging the use of satellite and the evolving state-of-the-art computing techniques. This has opened up a new window to reliably estimate and develop the hydroclimatic series for a longer period especially in the areas with scarce to no data. The GPPs find their way into hydroclimatic applications owing to their easy access and as cost-free resources. However, there is a plethora of GPPs (Darand & Khandu 2020) to choose from and no one-size-fits-all option available for selection. In addition, performance of GPPs is also not uniform across the spatial scale (Camici et al. 2018). They further possess a high level of uncertainty owing to their gridding algorithms, reliability and number of data sources used, kind of models used for estimating precipitation, gauge adjustment techniques, etc. (Aliyar et al. 2021; Ebert et al. 2007) that propagates through the application (hydrological modelling or extremes analysis). Thus, credibility of GPPs is subjective to factors like time, region, climatic zone, and the technology and methods involved in their development. Hence, evaluation of GPPs is crucial prior to their use. Evaluation and ranking techniques span across a multifaceted spectrum including the use of single criterion like (correlation coefficient (r) or root mean squared error (RMSE)) to multi-criteria decision making (MCDM) like compromise programming (Komaragiri & Kumar 2014; Zeleny 1973), Cooperative Game Theory (Gershon & Duckstein 1983), Preference Ranking Organization Method of Enrichment Evaluation (PROMETHEE-2) (Brans et al. 1986), and Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) (Opricovic & Tzeng 2004). These methods are gradually evolving and regularly improved/updated. Zeleny (1973) introduced compromise programming (CP) to automatically devise the solution closest to the reference dataset. This approach tends to obtain the best solution by finding the optimum values of evaluation indices under consideration. The search for the optimum value is guided by the entropy method to weigh each indicator (Komaragiri & Kumar 2014). The main advantage of the entropy method is

that it calculates the weight for each indicator without intervention from the evaluator avoiding unintentional bias. Further, the method weighs each indicator based on the information available which is given by its entropy value. Higher the entropy value, higher is the uncertainty.

Saying that, the primary purpose of the evaluation should be to select “fit-to-purpose” dataset rather than finding the “most correct” dataset. Abundance of studies on the evaluation of GPPs against the gauge-based station data is available that spans across a wide assortment of spatial coverage from global scale (Hobeichi et al. 2020; Shen et al. 2020; Yong et al. 2015), continental scale (Awange et al. 2016; Tarek et al. 2021), national scale (Aliyar et al. 2021; Darand & Khandu 2020; Prakash et al. 2016), regional scale (Ahmed et al. 2019; Camici et al. 2018; Satgé et al. 2020), and to basin scale (Duan et al. 2016; Fallah et al. 2020; Upadhyay et al. 2022; Yan Yang et al. 2014). Depending on region, scale, and evaluation method, the performance of GPPs is different. Lately, studies on the evaluation of GPPs across the Mekong River Basin (MRB), a transboundary river basin in Southeast Asia that transcends six countries China, Laos, Myanmar, Thailand, Cambodia, and Vietnam inhabited by more than 60 million people who rely on agriculture and fisheries (Piesse 2016), is growing. These studies can broadly be categorized into (a) ability to characterize the spatio-temporal distribution of precipitation (A. Chen et al. 2018; C.-J. Chen et al. 2017; Dandridge et al. 2019; Irannezhad & Liu 2022) and (b) blend of spatio-temporal characterization and ability to simulate hydrology (Dang Dinh et al. 2020; Tang et al. 2019, 2021; Tian et al. 2021; Try et al. 2020). Studies in former category mostly used relative bias (Bias), r , RMSE, and different variations in bias and errors as continuous statistical indices while probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) under categorical indices to evaluate the performance of GPPs. Similarly, the studies in the latter category used Nash–Sutcliffe efficiency (NSE), coefficient of determination (R^2), and percentage bias to evaluate hydrological simulation skill in addition to indices enlisted for category A.

All evaluation technique used in former studies focused on evaluating the GPPs based on their ability to reproduce the mean precipitation and skill to simulate the discharge in the basin. None of these studies considered the ability of GPPs to reproduce the extreme indices. The studies in category A were subjective to the evaluation skill of the researcher as the conclusion was drawn based on institution of the researcher without considering the tradeoffs among the evaluation indices used. As the precipitation is inherent with the intrinsic errors of missing data, instrumentation errors, instrument relocation, observation errors, etc. It is often difficult to rank the dataset only based on the statistical indicators as different statistical indices may suggest different GPPs. Thus, this indicates the need for an approach that seeks the tradeoffs among the indicators by automatically weighing them and ranking

accordingly. For the studies in category B, the uncertainty in the GPPs to simulate hydrology is also compromised by altering the other water balance components of the hydrological models like actual evapotranspiration, infiltration, etc. (Tian et al., 2021). Further, they were mostly carried out considering the entire MRB and recommended Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE) for the entire basin. Also, gauged-based reference data used were mostly concentrated in the upper reaches of the basin (China) and Thailand and relatively less number of observed stations were considered for Laos, Cambodia and Vietnam (Dang Dinh et al. 2020; Tian et al. 2021). For the basin covering a wide assortment of terrain from rugged mountainous to flat terrain, the performance of GPPs is not uniform (Alexander et al. 2020).

Hence, overall objective of this study is to evaluate eight GPPs, namely, APHRODITE, Climate Hazard Group Infra-Red Precipitation with Station (CHIRPS), Climate Prediction Center (CPC), CPC MORPHing (CMORPH), Climate Research Unit (CRU), Southeast Asian OBServed (SAOBS), Tropical Rainfall Measuring Mission (TRMM), and Global Precipitation Climatology Center (GPCC) for Nam Ngum River Basin (NRB) and Vietnam Mekong Delta (VMD), two sub-basins of LMRB featuring different characteristics in terms of climate and terrain. The specific objectives are to (i) evaluate GPPs using automated entropy-compromise programming (ECP) that automatically assigns weight to evaluation indicators and rank them, (ii) consider six extreme climate indices apart from mean precipitation, and (iii) check where a single GPP can be generalized for two sub-basins as indicated by former studies.

2 Study area

The study area comprises of two sub-basins of the LMRB: Nam Ngum River Basin in Lao PDR and Mekong delta in Vietnam as presented in Fig. 1

2.1 Nam Ngum River Basin

The NRB is a one of the important Mekong tributary sub-basins in central Lao PDR housing an area of 16,800 km² accounting for 7% of the country's area. The basin lies approximately between 17.9° N to 19.8° N and 101.85° E to 103.5° E. The main tributary of the basin, Nam Ngum River, traverses 420 km southwards before draining into Mekong River. It further discharges 40% of the country's flow to Mekong River which accounts for 14% of its flow (Meema et al. 2021). The basin is characterized by relatively flat in the southern part and steep topography with undulating terrain in the northeast part. The elevation for the Nam Ngum ranges between 151 and 2698 m.

The basin is marked by the tropical climate with distinct wet (June to September) and dry (October to April) seasons. March and April months usually witness the highest temperature with the average temperature ranging between 30 and 38 °C depending on the altitude and location. The average annual precipitation across basin ranges between 1450 and 3500 mm with basin average of 2000 mm (Meema et al. 2021).

The basin is significant as it has tremendous potential for hydropower and already houses four dams with Nam Ngum 1 being the largest one and many other are under construction or planned. The mean annual river flow of the NRB is approximately 21,000 Mm³ which attains its lowest level during March–April and reaches the peak during August–September.

2.2 Vietnam Mekong Delta

The VMD lies approximately between 8.55°N to 11.05°N and 104.45°E to 106.85°E in the southern Vietnam bounded by the East Sea in the south and southeast, Cambodia in the north, and Gulf of Thailand in the southwest. The basin is spread across an area of 40,000 km² and is drained by a complex network of the canals, rivers, and dikes.

The VMD is characterized by the relatively flat terrain < 5 m. It observes tropical climate with distinct wet (May to November) and dry season (December to April). About 30–40% of the delta is inundated during rainy (wet) season influenced by the Indian Summer Monsoon (Duy et al. 2021). The average annual precipitation across the basin ranges between 1300 and 2500 mm of which 80–90% is attributed by the precipitation during the wet season accounting for a discharge load of 75–85% (Le et al. 2021).

The VMD is well known for the aquaculture and rice production occupying about 50% (~ 1.9 million ha) of the agricultural land (Sebesvari et al. 2012). However, the annual floods from the Mekong River inundates the large parts of the VMD.

3 Datasets

3.1 Observation datasets

The observed gauge-based precipitation dataset in the NRB were very limited. Only six stations data were made available by the Department of Meteorology and Hydrology, Laos that were mostly concentrated around the central region of the Basin along the Nam Ngum River. The spatial distribution of the observed dataset is presented in Fig. 1. The daily observed gauge-based precipitation dataset was available for the reference period of 1998 to 2014.

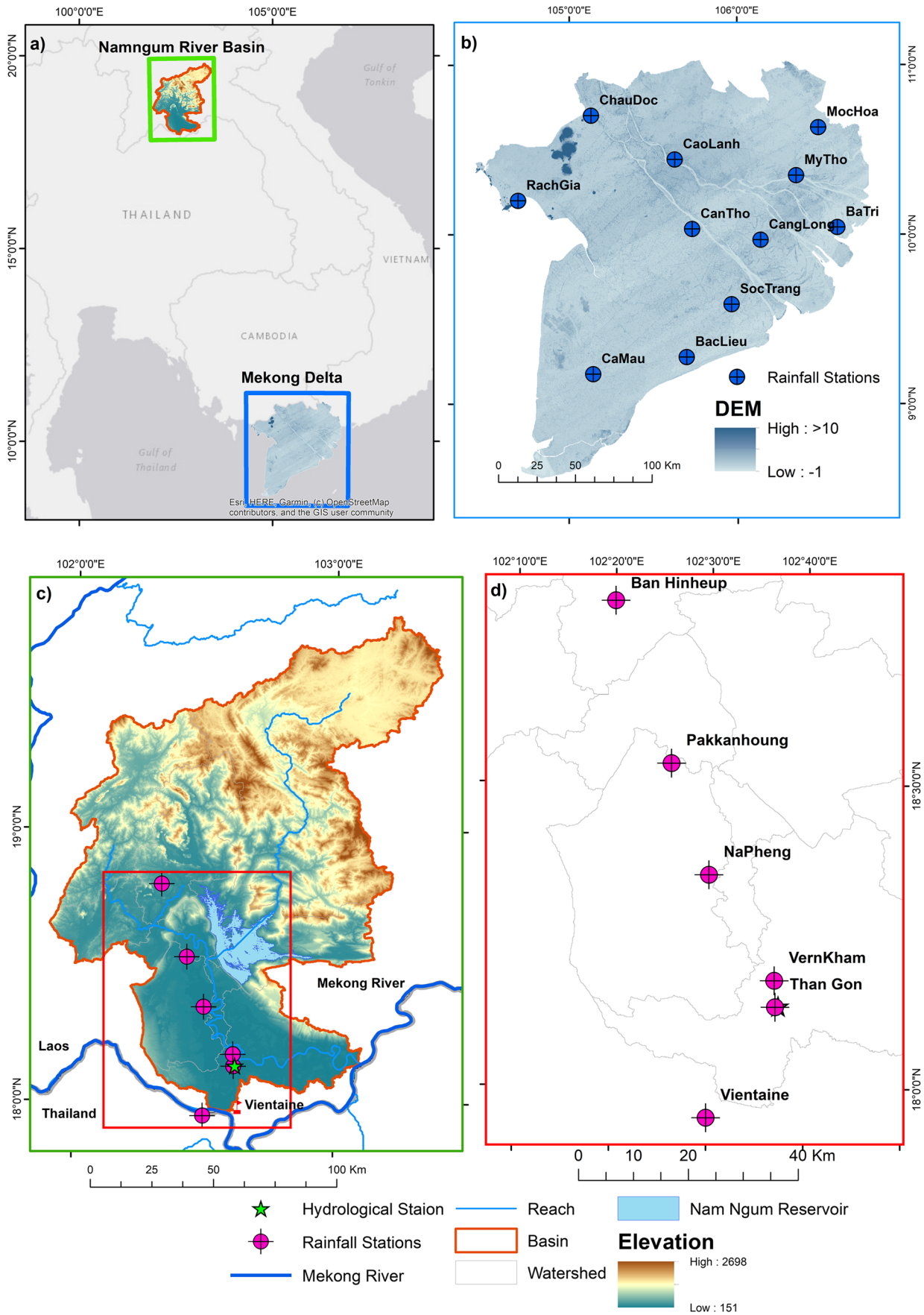


Fig. 1 **a** Location map of Nam Ngum River Basin and Vietnam Mekong Delta. **b** Elevation profile along with the spatial distribution of reference precipitation gauges across the Vietnam Mekong Delta. **c** Elevation profile of Nam Ngum River Basin. **d** Spatial distribution of the reference precipitation gauges in the Nam Ngum River Basin

Only five stations having the complete observed data for the overlapping timeframe were considered for NRB. In case of the VMD, eleven rain gauge stations were made available by the Hydro-Meteorology Data Center, Hanoi, Vietnam. Since, the availability of the gauge density is low in both basins and is unlikely to represent the spatial variability in the rainfall, the GPPs are considered for evaluation. The sparse gauge-based dataset is not well distributed to be considered for the basin scale applications as the uncertainty associated with the sparsity is likely to propagate in the application like hydrological modelling or trends or extremes, etc.

3.2 Gridded precipitation products

In this study, we evaluated eight GPPs that comprise five gauge-based, one satellite-based, and two reanalyzed precipitation products with different spatial and temporal resolutions. Different GPPs have different temporal resolution, and an overlapping reference period of 1998–2014 was considered for the study. Details of products, their description, and source from where they can be retrieved are presented in Table 1.

3.2.1 Gauge-based GPPs

APHRODITE Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE) is a observed gauge-based interpolated precipitation product using improved angular-distance-weighting interpolation technique that covers three spatial domains (Monsoon Asia, Middle East, and Russia) at the spatial resolution of 0.25° and of daily temporal resolution (Yatagai et al. 2012). The Meteorological Research Institute of Japan Meteorological Agency and the Research Institute for Humanity and Nature are the developers of the dataset. Daily based precipitation records form a wide assortment of sources including the global telemetric stations, individual researcher's collections, meteorological stations records of the National Climatic Data Center, etc. were obtained to develop the APHRODITE. Two versions of products are available: V1101 is available from 1951–2007 and V1101EX_R1 from 2007–2015).

CPC The Climate Prediction Center (CPC) precipitation product is a quality-controlled observed gauge-based interpolated daily scale dataset produced by the National Oceanic

and Atmospheric Administration (NOAA) at the spatial resolution of 0.5° covering a period of 1979 until present (Xie et al. 2007). The data sources used in developing these products house more than 30,000 gauge station data provided by Cooperative Observer Network (COOP), GTS, and numerous national and international organizations.

CRU The Climate Research Unit (CRU) is a meteorological stations based interpolated terrestrial only (except Antarctica) monthly dataset at a spatial resolution of 0.5° covering a period of 1901 to 2020 (Harris et al. 2014; Ian Harris et al. 2020). The anomalies for the gauge-based precipitation data were gridded using the angular distance weighting (ADW) interpolation algorithm and were converted to the actual values of precipitation (not anomalies). The dataset is developed by the University of East Anglia with a support from Natural Environment Research Council (NERC), United Kingdom (UK), the United States (US) Department of Energy, and the UK National Center for Atmospheric Sciences (NCAS). In this study, the CRU gridded Time Series version 4.05 (CRU TS v4.05) was used.

GPCC The Global Precipitation Climatology Center (GPCC) is a gauge-based interpolated long-term gridded monthly dataset with a spatial resolution of 0.25° covering a period of 1891–2016. The dataset houses about 50,000 stations per month throughout the globe (Zandler et al. 2019). The raw gauge-based dataset to produce GPCC was supported by several institutions including national meteorological institutions (NMAs), the World Meteorological Organization (WMO), the Food and Agricultural Organization (FAO), the Climate Research Unit (CRU), etc.

SAOBS The Southeast Asian OBServed (SAOBS), a product of Southeast Asia Climate Assessment and Dataset (SACAD), is an observed gauge-based interpolated daily dataset for Southeast Asia at two spatial resolution of 0.25° and 0.5° from 1981 to 2014 (van den Besselaar et al. 2017). About 4000 gauge-based precipitation datasets from Southeast Asian countries were used to develop the SAOBS. This study used the dataset at 0.25° spatial resolution.

3.2.2 Satellite-based GPP

TRMM The Tropical Rainfall Measuring Mission (TRMM) is a satellite-based precipitation estimates developed by a joint mission between the National Space Development Agency (NASDA), Japan and the US National Aeronautics and Space Administration (NASA) with an aim to estimate rainfall in the tropical and subtropical regions. This daily dataset is available at the spatial resolution of 0.25° between 1998 until 2020 (Huffman & Bolvin 2018).

Table 1 Different GPPs categorized according to their type they fit in, their spatial and temporal resolutions, period of their availability, and sources from where they can be retrieved

Type	Products	Spatial resolution	Temporal resolution	Spatial coverage	Period of record	Data source
Gauge-based	APHRODITE (V1101 and V1101EX_R1)	0.25°	Daily	Monsoon Asia, Russia and Middle East	1951–2015	Data can be downloaded from: http://aphrodite.st.hirosaki-u.ac.jp/download/ (user registration is required)
	CPC	0.5°	Daily	Global	1979–present	Data can be downloaded from: https://www.esrl.noaa.gov/psd/data/gridded/Tables/precipitation.html
	CRU TS4.05	0.5°	Monthly	Global	1901–2015	Data can be downloaded from: https://data.ceda.ac.uk/badc/cru/data/cru_ts/cru_ts_4.05/data (registration required to access data)
	GPCC	0.25°	Monthly	Global	1891–2016	Data can be downloaded from: https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-monthly_v2018_doi_download.html
	SAOBS	0.25°,0.5°	Daily	Southeast Asia	1981–2014	Data can be downloaded from: https://sacad.database.bmkg.go.id/download/grid/download.php
Reanalyzed	CHIRPS25	0.25°	Daily	50°N–50°S	1981–present	Data can be downloaded from: https://data.chc.ucsb.edu/products/CHIRPS-2.0/
	CMORPH	0.25°,0.5°	Daily	60°N–60°S	1998–present	Data can be downloaded from: https://www.ncei.noaa.gov/data/cmorph-high-resolution-global-precipitation-estimates/access/daily/
Satellite-based	TRMM 3B42_7	0.25°	Daily	50°N–50°S	1998–present	Data can be downloaded from: https://disc.gsfc.nasa.gov/datasets/TRMM_3B42_Daily_7/summary?keywords=TRMM_3B42

3.2.3 Reanalyzed GPPs

CHIRPS The CHIRPS is a reanalyzed precipitation product developed from the blend of multiple sources of dataset including observed rain gauge data, precipitation based on cold cloud duration (CCD) infrared data of National Oceanic and Atmospheric Administration (NOAA)—National Climate Data Center (NCDC), TRMM 3B42 and Version 2 atmospheric model rainfall of NOAA-Climate Forecast System (CFS). The dataset is available at two spatial resolutions (0.05° and 0.25°). The CHIRPS25 is the daily precipitation product at the spatial resolution of 0.25° spanning across a period of 1981 until present (Funk et al. 2015) which was used in this study.

CMORPH The NOAA CPC MORPHing (CMORPH) dataset is a bias-corrected precipitation data product developed by blending gauge-based observations with passive microwave-based infrared satellite precipitation estimates using CPC Morphing Technique. CMORPH data is available at three spatial resolutions of 8 km (only December 2002 onwards), 0.25° and 0.5° and three temporal resolution of 30 min,

3-hourly and daily from 1998 until present (Xie et al. 2017). This study used the CMORPH with 0.25° spatial resolution and daily temporal resolution.

4 Methodology

The series of steps adopted for evaluation and ranking of the gridded dataset is illustrated in Fig. 2. It starts with the collection of observed and gridded datasets. The methodology adopts the pixel-to-point comparison (considering the limited number of gauge stations within the basin) for which the gridded datasets were extracted to station's coordinates. Following the data extraction, extreme indices were computed and homogeneity tests were carried out. Six extreme indices proposed by Experts Team on Climate Change Detection Indices (ETCCDI) of World Meteorological Organization (WMO) were computed for each dataset. List of extreme indices used in this study are presented in Table 2. A payoff matrix was then formulated by normalizing the evaluation metrics obtained for each EI in addition

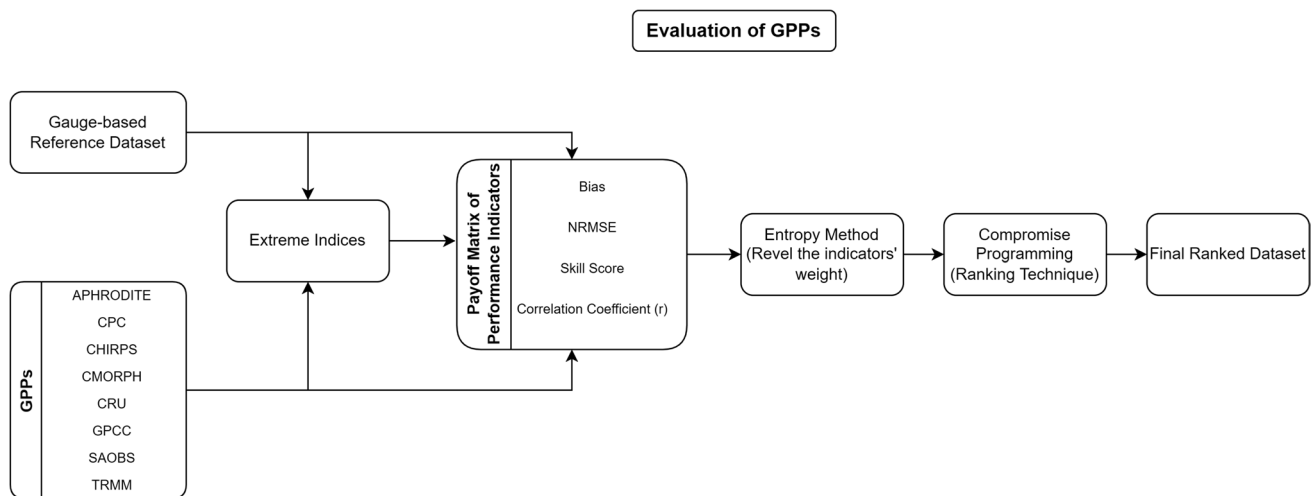


Fig. 2 Flowchart displaying the GPP evaluation and ranking techniques adopted in this study

Table 2 ETCCDI extreme precipitation indices

Index	Descriptive name	Definition	Unit
CDD	Consecutive dry days	Maximum number of consecutive dry days when precipitation is less than a specified threshold (precipitation < 1 mm)	Days
CWD	Consecutive wet days	Maximum number of consecutive dry days when precipitation is above a specified threshold (precipitation > 1 mm)	Days
R10mm	Heavy precipitation days	Annual count of days when daily rainfall rate (RR) > 10 mm	Days
R20mm	Very heavy precipitation days	Annual count of days when daily rainfall rate (RR) > 20 mm	Days
RX1day	Maximum 1-day precipitation amount	Annual maximum 1-day precipitation	mm
RX5day	Maximum 5-day precipitation amount	Annual maximum 5-day precipitation	mm

to mean precipitation for each station. Metrics normalization was carried out to avoid the dominance of the larger value indicators. Entropy method (Komaragiri & Kumar 2014) was employed to weigh each evaluation criteria. The payoff matrix along with the weights obtained from entropy method were then supplied to the compromise programming to obtain station wise ranked dataset. To obtain one final ranked dataset, sum of ranks for individual indices were computed and the one with the least value was ranked 1st, second least value was ranked 2nd, and so on.

4.1 Extreme indices

The Expert Team on Climate Change Detection and Indices (ETCCDI) of the WMO has developed 27 core extreme climate indices to analyze and standardize the climate extremes (Zhang et al. 2011). Of proposed 27 indices, six indices corresponding to precipitation extreme analysis used in this study are presented in the Table 2.

4.2 Continuous statistical indices

To evaluate the GPPs against the available observed dataset, the following continuous statistical indices were used:

Pearson correlation coefficient (r)

$$r = \frac{\sum_{i=1}^n (o_i - \bar{o})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (o_i - \bar{o})^2} \sqrt{\sum_{i=1}^n (p_i - \bar{p})^2}} \quad (1)$$

where o_i , \bar{o} , p_i , \bar{p} , and n are observed precipitation for i^{th} position, observed mean for the studied period, gridded precipitation for i^{th} position, average gridded precipitation data for the studied period, and n is the total number of data in the timeseries. Pearson correlation coefficient measures the degree of similarity between the observed and gridded dataset. Its value ranges between -1 and 1 . Value of 1 (-1) represent a perfect positive (negative) correlation.

Bias

$$\text{Bias} = \sum_{i=1}^n p_i - o_i \quad (2)$$

where p_i and o_i are gridded and observation precipitation data at i^{th} position and n is the total number of datapoints in time series. The positive (negative) value of Bias indicates overestimation (underestimation) by the gridded dataset while the value of zero shows perfect estimation. The closer the value of Bias to zero better is the estimation.

Root mean squared error

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (3)$$

where x_i, y_i are the observed and gridded data points at i^{th} position and n is the total number of datapoints in the timeseries. RMSE measures the differences between the observed and gridded dataset and estimates the averaged error magnitude. Value of RMSE close to zero indicate the better agreement between gridded and observed dataset.

5 Skill score

The skill score compares the probability density function (pdf) of the GPPs with the pdf of the observed/reference dataset and measure the common area between the pdfs.

$$S_{\text{score}} = \sum_i^N [\min(P_i^{\text{GPP}}, P_i^{\text{reference}})] \quad (4)$$

where S_{score} is the skill score, N is the number of bins to calculate pdf, P_i^{GPP} and $P_i^{\text{reference}}$ are the frequencies of values in a bin of the GPP and observed dataset, respectively. A skill score of 1 indicate a complete overlap between the pdfs of GPP and observed and a value of 0 indicate no overlap at all.

5.1 Performance evaluation based on categorical indices

Apart from the continuous statistical indices, three categorical indices—probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI)—were used to evaluate the skills of GPPs in accurately predicting the rain/no rain events. POD measures the ability of the GPPs to correctly estimate the precipitation of a given threshold. It is also commonly known as the hit ratio. Similarly, FAR measures the number of false precipitation event GPPs detected that were not present in the observed data. Likewise, the CSI measures the ability of GPPs to estimate the precipitation skill. The POD, FAR, and CSI are calculated using Eqs. 5, 6, and 7.

$$\text{POD} = \frac{R}{R + M} \quad (5)$$

$$\text{FAR} = \frac{F}{R + F} \quad (6)$$

$$\text{CSI} = \frac{R}{R + F + M} \quad (7)$$

where R is the number of hits, F is the number of false alarm, and M is the number of misses. The POD, FAR, and CSI ranges between 0 and 1 , where 1 indicates perfect skill in detecting the precipitation event for POD and CSI. In

contrast, 0 represents the perfect score for FAR, indicating that GPPs did not give any false alarm.

5.2 Entropy method

Entropy method assigns weights to each evaluation criteria under study. Equations 8–11 are used to compute weights using entropy method.

$$En_j = -\frac{1}{\ln(T)} \sum_{a=1}^T k_{aj} \ln(k_{aj}) \text{ for } j = 1, 2, \dots, J \tag{8}$$

$$k_{aj} = \frac{k_j(a)}{\sum_{a=1}^T k_j(a)} \tag{9}$$

where En_j is the entropy for each evaluation criteria j , T is the number of gridded precipitation datasets, and k_{aj} is the value of the j^{th} indicator for a^{th} gridded dataset.

The entropy (En_j) is then used to compute the degree of diversification and weights for each index as given by following equations.

$$Dd_j = 1 - En_j \tag{10}$$

$$r_j = \frac{Dd_j}{\sum_{j=1}^J Dd_j} \tag{11}$$

where Dd_j and r_j represent the degree of diversification and normalized weight of indicator j , respectively.

The higher the value of entropy, the higher is the uncertainty associated with the indicator and less is the weight.

5.3 Compromise programming

Compromise programming is the ranking technique that utilizes minimum distance of the indicator associated with GPP and the observed dataset. The distance formula for compromise programming is given by Eq. 12.

$$L_p(a) = \left[\sum_{j=1}^J w_j |f_j^* - f_j(a)|^p \right]^{\frac{1}{p}} \tag{12}$$

where $L_p(a)$ is the minimum distance metric for the a gridded dataset, j is the indicator, f_j^* and $f_j(a)$ are normalized value of j^{th} indicator for observed and a^{th} gridded dataset, respectively, w_j is the weight of j^{th} indicator obtained from entropy method, and p represents distance parameter ($p = 1$ for linear and $p = 2$ squared Euclidean distance). The lower the value L_p places the gridded dataset on the top and so on.

5.4 Homogeneity test

To test the inhomogeneities possibly present in dataset that are likely to influence the evaluation process, three of the four tests proposed by Wijngaard et al. (2003), which are standard homogeneity test, Pettitt test, and Buishand test, were considered for this study. As few datasets (GPCC and CRU) lack daily precipitation, annual total precipitation was used over the annual wet days (precipitation > 1 mm) for the homogeneity test. Similar approach adopted by Zandler et al. (2019) was used. The pyHomogeneity package in python 3.7 environment was used to test homogeneity of datasets at the significance level of 1% for the entire study period of 1998 to 2014. The obtained test results from these tests were then condensed into three categories: (a) “useful” if at most one test indicated inhomogeneous dataset, (b) “doubtful” if two tests indicated inhomogeneity, and (c) “suspect” otherwise.

6 Results

6.1 Homogeneity tests’ result

The homogeneity test indicated that majority of GPPs (> 90%) were homogeneous for both NRB and VMD. In case of NRB, only CPC was doubtful for three stations, namely, Ban Hinheup, Naphok, and Pakkanhoung. The condensed categories of the homogeneity tests for NRB is presented in Table 3.

U useful, *D* doubtful.

Table 3 Categories of the homogeneity tests for different GPPs against each station for NRB

	BanHinheup	Napheng	Naphok	Pakkanhoung	Vientaine
APHRODITE	U	U	U	U	U
CHIRPS	U	U	U	U	U
CMORPH	U	U	U	U	U
CPC	<i>D</i>	U	<i>D</i>	<i>D</i>	U
CRU	U	U	U	U	U
GPCC	U	U	U	U	U
SAOBS	U	U	U	U	U
TRMM	U	U	U	U	U

The italicized letter indicate the inhomogeneity present in data set as suggested by two or more homogeneity tests.

On the other hand, three GPPs indicated signs of inhomogeneities for some stations in VMD. For instance, CRU was doubtful for BaTri station and CPC was doubtful for two stations and indicated suspect for BacLieu station while SAOBS was doubtful for 10 stations except Chau-Doc. Results of homogeneity test for NRB and VMD are presented Table 4.

The p-stat obtained for individual test for all stations to generate the condensed categories of homogeneity analysis is presented in Appendix Tables (8) and (9), respectively.

6.2 Performance of GPPs across NRB based on continuous statistical indices and CP

The comparison of the GPPs with the station data were made on a grid-to-point basis whereby nearest grid data were extracted to the station as enough stations were unavailable to cover the entire basin. Station-wise evaluation indicators (r , Bias, RMSE, and SS) are presented in Table 5 and Fig. 3 for NRB. The APHRODITE in daily scale and CRU and GPCC at monthly time resolution showed a good fit against the observed dataset with correlation coefficient ranging between 0.85 and 0.97. The GPCC, TRMM, and APHRODITE have least bias among all the GPPs. The former two GPP represent positive bias ranging between 1.21 and 38 mm/month, indicating overestimation of precipitation while APHRODITE underestimated the precipitation for 80% of stations (-4.97 to -25.34 mm/month). Rest all GPPs (except CHIRPS) highly underestimated the precipitation in the NRB. CPC accounts for the highest negative bias in monthly precipitation for all stations followed by the CMORPH. The APHRODITE performed best in terms of RMSE with its values ranging between 45.32 and 111.55 mm/month followed by TRMM (48.83 to 109.87 mm/month) and GPCC (34.73 to 126.37 mm/month). Rest all GPPs were relatively erroneous.

The skill score (SS) represents the coincidence of two PDFs. The SS of APHRODITE and TRMM ranged between 0.69 to 0.90 outperformed the other datasets. The CPC,

CMOPRH, and SAOBS exhibited larger differences between the PDFs against observed dataset. The worst was exhibited by the CPC with SS value ranging between -0.12 and 0.4 (except for Vientiane). Rest GPPs exhibited moderate SS. The discrepancies in the evaluation indices between GPPs can also be attributed to the different interpolation techniques and the station density available for interpolation (Hu et al. 2018).

The Taylor diagram (Taylor 2001) in Fig. 3 visually illustrates the station-wise performance of different gridded dataset against the observed dataset with respect to three different evaluation indicators (r , RMSE, and SD), and a table in the middle shows the final rank of the dataset. The black secants represent the SD while the green secants are for RMSE and the dotted blue straight lines represent r . The red secant signifies the SD of the reference dataset. In terms of SD, GPCC and TRMM were close to observed SD than other GPPs. Before obtaining the final rank, datasets with daily temporal resolution were further evaluated for the extreme climate indices. The performance of individual dataset against the observed dataset in capturing the variability of EIs is presented in Appendix Fig. 7. As monthly datasets were evaluated for the mean precipitation only, their score summed up to minimum and placed GPCC and CRU in the top order. It can be seen from the Fig. 3 and Table 5 that APHRODITE secured 3rd rank (overall) and 1st (at daily scale) for about 60% stations of NRB.

Figure 4 houses the ranking of each GPP against the observed dataset in the form of heatmaps obtained from entropy method and compromise programming while Appendix Fig. 8 presents the station-wise box and whisker plot of EIs for the different GPPs with detailed description. Varying degree of performances were observed for GPPs for different stations. For instance, SAOBS was able to capture the CDD for 60% of stations followed by CPC. On the other hand, CPC and CHIRPS did well in capturing the CWD for 80% of the stations combined. APHRODITE and CMORPH were able to capture the range of CDD and CWD, but they

Table 4 Categories of condensed homogeneity tests for GPPs against each station for VMD

	Bac Lieu	Ba Tri	Ca Mau	Cang Long	Can Tho	Cao Lanh	Chau Doc	Moc Hoa	My Tho	Soc Trang	Rach Gia
APHRODITE	U	U	U	U	U	U	U	U	U	U	U
CHIRPS	U	U	U	U	U	U	U	U	U	U	U
CMORPH	U	U	U	U	U	U	U	U	U	U	U
CPC	<i>S</i>	U	<i>D</i>	U	U	U	U	U	U	U	<i>D</i>
CRU	U	<i>D</i>	U	U	U	U	U	U	U	U	U
GPCC	U	U	U	U	U	U	U	U	U	U	U
SAOBS	<i>D</i>	<i>D</i>	U	<i>D</i>	<i>D</i>	<i>D</i>	U	<i>D</i>	<i>D</i>	<i>D</i>	<i>D</i>
TRMM	U	U	U	U	U	U	U	U	U	U	U

U useful, D doubtful, S suspect.

The italicized letter indicate the inhomogeneity present in data set as suggested by two or more homogeneity tests.

Table 5 Statistical measures of GPPs for different statistical indices across stations of NRB

		BIAS	RMSE	<i>r</i>	SS
APHRODITE	Ban_Hinheup	25.53	97.53	0.88	0.73
	Napheng	-25.34	109.96	0.90	0.76
	Naphok	-4.97	73.93	0.88	0.77
	Pakkanhoung	-22.00	111.55	0.88	0.73
	Vientiane	-13.06	45.32	0.96	0.90
CHIRPS	Ban_Hinheup	100.97	175.26	0.86	0.14
	Napheng	0.95	107.30	0.88	0.77
	Naphok	33.22	93.91	0.86	0.63
	Pakkanhoung	55.08	135.59	0.86	0.61
	Vientiane	-3.89	53.67	0.93	0.86
CRU	Ban_Hinheup	-43.25	113.06	0.85	0.64
	Napheng	-45.10	125.52	0.89	0.68
	Naphok	-8.96	80.85	0.85	0.73
	Pakkanhoung	-49.78	135.95	0.85	0.60
	Vientiane	-5.26	53.12	0.93	0.86
CPC	Ban_Hinheup	-116.82	200.11	0.52	-0.12
	Napheng	-101.02	184.96	0.80	0.31
	Naphok	-65.03	119.28	0.79	0.40
	Pakkanhoung	-123.35	211.93	0.68	0.04
	Vientiane	-20.61	60.63	0.92	0.82
GPCC	Ban_Hinheup	38.28	126.37	0.83	0.55
	Napheng	5.11	111.82	0.87	0.75
	Naphok	19.75	77.48	0.89	0.75
	Pakkanhoung	1.21	118.07	0.84	0.70
	Vientiane	6.54	34.73	0.97	0.94
SAOBS	Ban_Hinheup	-43.61	152.80	0.66	0.35
	Napheng	-60.74	147.85	0.81	0.56
	Naphok	-24.18	99.31	0.79	0.59
	Pakkanhoung	-52.78	144.96	0.78	0.55
	Vientiane	-22.51	82.32	0.84	0.67
CMORPH	Ban_Hinheup	-74.16	142.60	0.81	0.43
	Napheng	-90.35	162.85	0.90	0.47
	Naphok	-40.75	95.49	0.84	0.62
	Pakkanhoung	-78.51	156.21	0.85	0.48
	Vientiane	-37.71	70.94	0.92	0.75
TRMM	Ban_Hinheup	14.48	101.13	0.86	0.71
	Napheng	10.24	101.58	0.89	0.79
	Naphok	31.42	85.66	0.89	0.69
	Pakkanhoung	12.94	109.87	0.86	0.74
	Vientiane	15.28	48.83	0.95	0.88

underestimated (overestimated) for CDD (CWD) for almost all the stations. For about 60% of the stations, SAOBS was able to represent the R10mm precipitation followed by APHRODITE. Similarly, in case of R20mm precipitation, APHRODITE ranked itself in top order for 60% of the stations. APHRODITE and TRMM were able to represent the Rx1day precipitation for 80% of the stations combined. In

contrast, SAOBS dominated the order for 60% of the stations for Rx5mm indices.

6.3 Evaluation GPPs based on categorical indices for NRB

The skill of individual GPPs to correctly discriminate the rain/no-rain events was further tested using three categorical indices (POD, FAR, and CSI) using 1 mm as the precipitation threshold. The values of POD, FAR, and CSI for each GPP is presented in Table 6. APHRODITE dataset had the highest POD ranging between 0.96 and 0.98 while SAOBS had the least FAR value ranging between 0.39 and 0.46. The TRMM had the highest CSI value ranging between 0.44 and 0.64.

6.4 Performance of GPPs across VMD based on continuous statistical indices and CP

The results for evaluation criteria of the different GPPs against the observed stations data for VMD is presented in Appendix Table (10). APHRODITE, CHIRPS, GPCC, and TRMM demonstrate least bias for almost all the stations ranging between ± 23 mm/month. The former two GPPs mostly have the negative bias and later two GPPs with the positive bias. Similarly, CMORPH, CPC, and SAOBS mostly underpredicted the precipitation as high as -184 mm/month for RachGia station (SAOBS). Correspondingly, in terms of RMSE, APHRODITE, GPCC, and TRMM has the least RMSE with values less than 86 mm/month followed by CHIRPS, CMORPH, CPC, and CRU. The SAOBS had poor RMSE ranging between 150 and 241 mm/month. Likewise, except for the SAOBS, all GPPs demonstrated a good correlation (above 0.7 for 80% of stations) with the observed dataset at the monthly scale with higher correlation corresponding to TRMM and GPCC. The *r* values for SAOBS ranged between 0.16 and 0.42 reveal very poor correlation at the monthly timescale. Figure 5 graphically illustrates the performance of each GPPs in terms of *r*, RMSE, and SD for VMD.

The GPCC had the better coincidence of PDF with the observed indicated by the SS values ranging between 0.52 and 0.98 for different stations. Rest all GPPs has the similar performances in terms of SS except (CRU and SAOBS). The SAOBS has the least overlapping PDF with observed as indicated by the negative values of SS (-1.28 to -0.82).

Similarly, Appendix Fig. 8 presents a station-wise box and whisker plot housing the performance of daily-scale GPPs against observed dataset for several climate EIs under study for VMD. During the study period from 1998 to 2014, TRMM outperformed all other GPPs in terms of representing the climate EIs against EIs of observed dataset for all the stations. Strong differences can be observed for other GPPs in representing the EIs. For

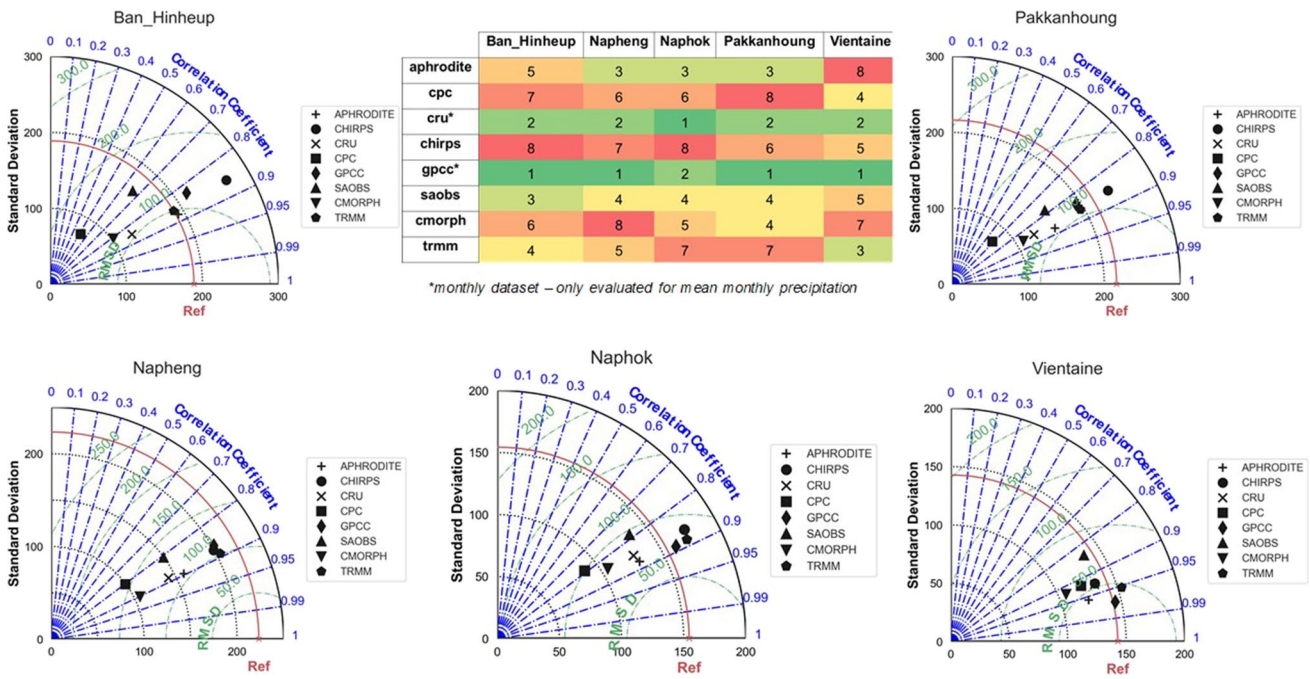
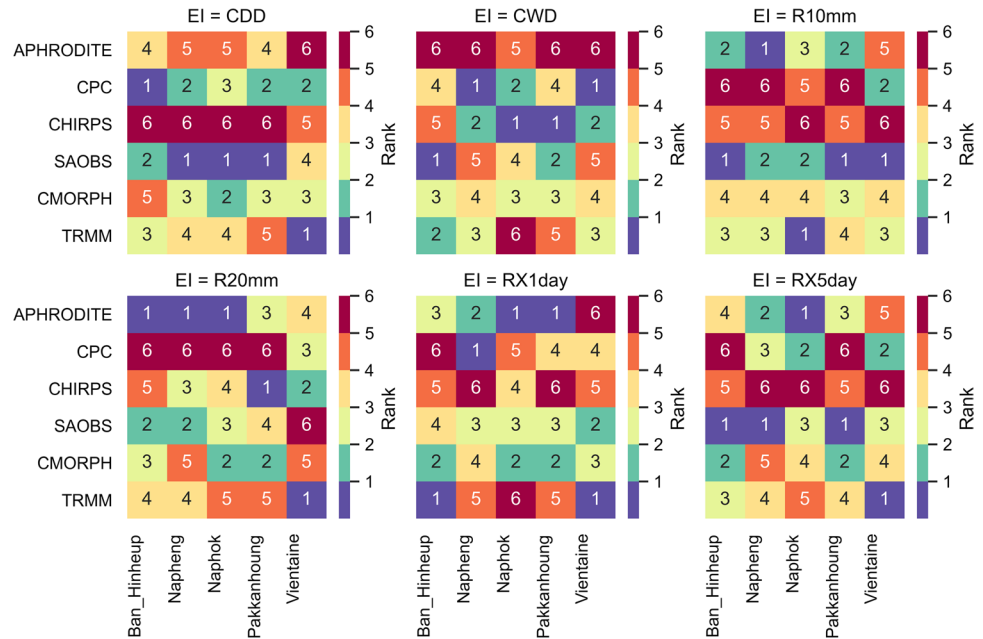


Fig. 3 Taylor diagram displaying performance of different datasets against observed dataset for different evaluation indicators

Fig. 4 Heatmaps capturing rank of GPPs evaluated against the observed dataset for different EIS for all the stations in NRB



instance, TRMM and CMORPH closely represent CDD (CWD) while other GPPs underestimate (overestimate) CDD(CWD) for almost all the stations. Similarly, TRMM and APHRODITE adequately represent the observed R10mm EI for almost all the stations while other GPPs underestimated R10mm EI except for CHIRPS which overestimated the R10mm EI for all stations. However,

only TRMM seems to represent R20mm while other GPPs underestimate it for all the stations. Likewise, only TRMM and CMORPH sufficiently captured the Rx1day and Rx5day indices while other GPPs underestimated them. Detail description of performance of GPPs in capturing the variability of EIs in observed dataset along with statistics is presented in Appendix Fig. 8.

Table 6 Evaluation of GPPs based on the categorical indices (bold values represent better performance for each index)

Datasets	Cis	Ban_Hinheup	Napheng	Naphok	Pakkanhoung	Vientiane
APHRODITE	POD	0.98	0.98	0.96	0.98	0.98
	FAR	0.53	0.53	0.58	0.54	0.42
	CSI	0.46	0.47	0.41	0.46	0.57
CHIRPS	POD	0.79	0.82	0.83	0.80	0.83
	FAR	0.47	0.45	0.51	0.49	0.33
	CSI	0.47	0.49	0.45	0.45	0.59
CMORPH	POD	0.91	0.89	0.88	0.87	0.88
	FAR	0.42	0.42	0.47	0.45	0.28
	CSI	0.55	0.54	0.49	0.51	0.65
CPC	POD	0.84	0.88	0.91	0.83	0.95
	FAR	0.46	0.45	0.50	0.48	0.33
	CSI	0.49	0.51	0.48	0.47	0.65
SAOBS	POD	0.81	0.80	0.83	0.82	0.83
	FAR	0.39	0.41	0.46	0.42	0.30
	CSI	0.53	0.52	0.48	0.51	0.61
TRMM	POD	0.87	0.86	0.82	0.83	0.81
	FAR	0.41	0.42	0.47	0.52	0.25
	CSI	0.54	0.53	0.48	0.44	0.64

CI's categorical indices.

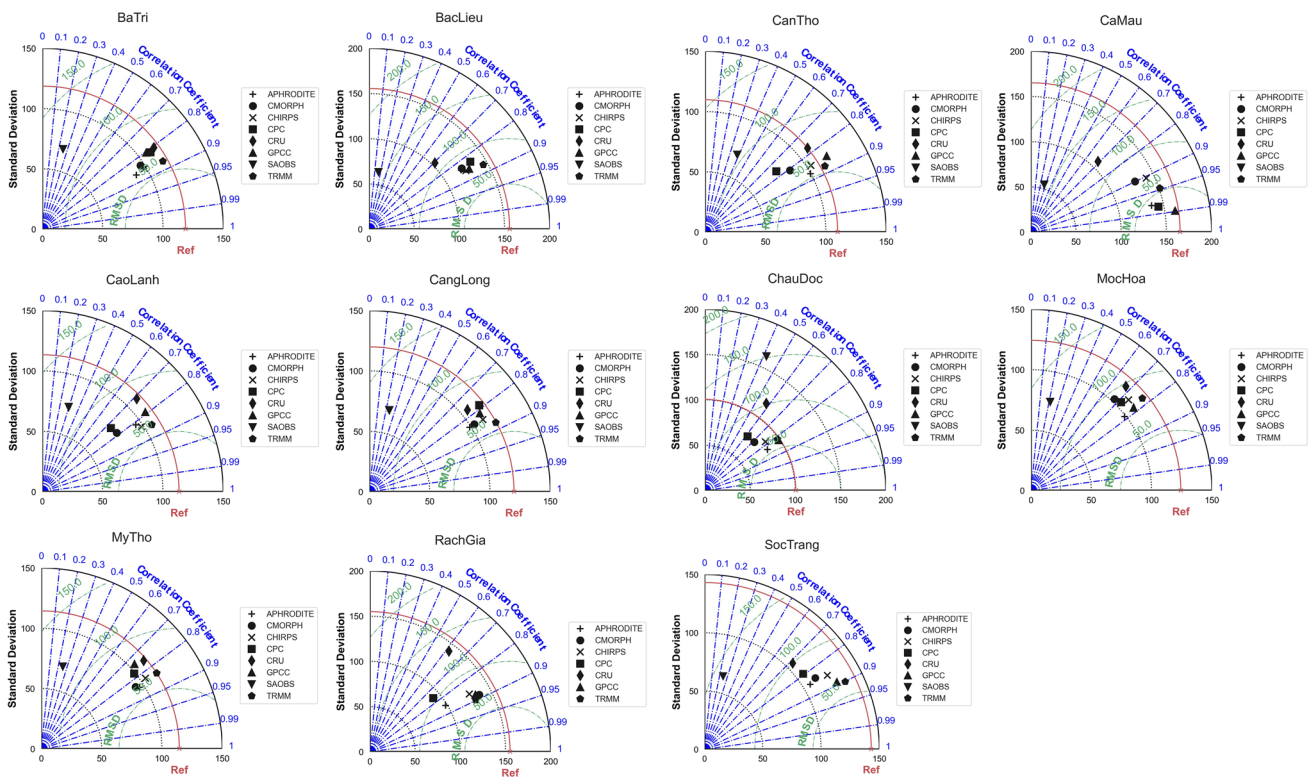
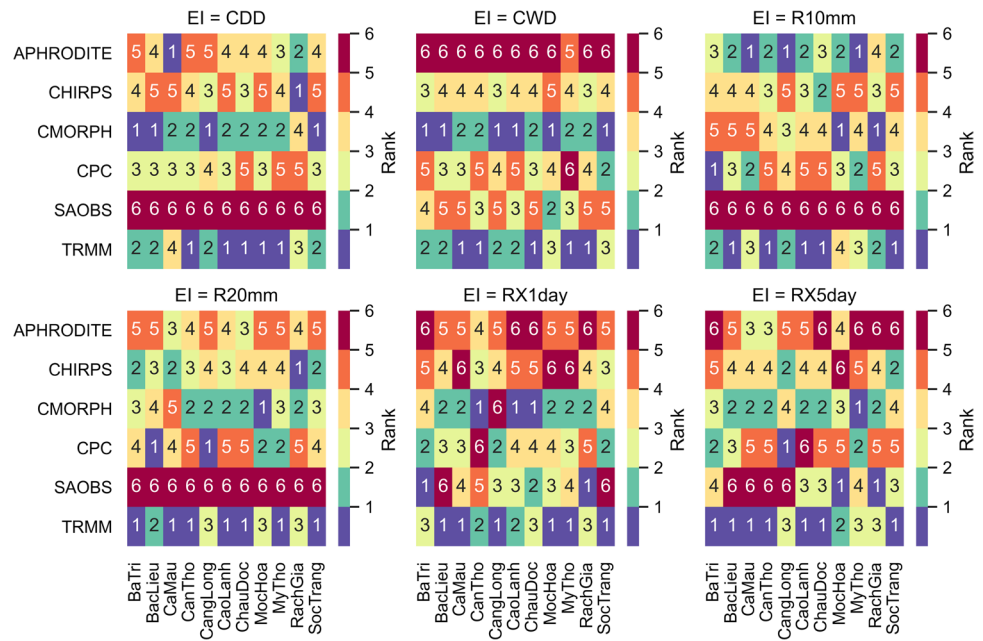


Fig. 5 Taylor diagram displaying performance evaluation of GPPs against observed dataset at different stations for Mekong Delta

Fig. 6 Heatmaps capturing rank of GPPs evaluated against the observed dataset for different EIS for all the stations in Mekong Delta



The rank of each GPP based on its performance against the observed dataset for different climate indices is presented by the heatmap in the Fig. 6.

6.5 Evaluation of GPPs based on categorical indices

The threshold to discriminate rain/no-rain was considered 1 mm as recommended by WMO. Based on the different categorical indices used to evaluate the GPPs, APHRODITE seems to

perform well in detecting the rainfall as indicated by the POD value greater than 0.85 for all stations as presented in Table 7. The APHRODITE also has the higher FAR values ranging between 0.34 and 0.52. This can be attributed to the higher number of CWD and fewer CDD for APHRODITE. Satellite-based estimate TRMM and reanalyzed dataset CMORPH have the least value for POD not above 0.71 for all stations. They also have relatively lower FAR compared to station-based gridded dataset. However, the overall critical index is similar for all the datasets.

Table 7 Evaluation of GPPs based on categorical indices for Mekong Delta

Datasets	CI	Bac Lieu	Ba Tri	Ca Mau	Cang Long	Can Tho	Cao Lanh	Chau Doc	Moc Hoa	My Tho	Soc Trang	Rach Gia
APHRODITE	POD	0.90	0.90	0.95	0.91	0.92	0.87	0.85	0.88	0.87	0.88	0.87
	FAR	0.45	0.50	0.34	0.48	0.47	0.51	0.51	0.49	0.52	0.42	0.45
	CSI	0.52	0.47	0.63	0.49	0.51	0.46	0.45	0.47	0.45	0.54	0.51
CMORPH	POD	0.70	0.66	0.70	0.70	0.63	0.58	0.55	0.61	0.66	0.69	0.71
	FAR	0.32	0.35	0.34	0.35	0.35	0.39	0.43	0.44	0.38	0.30	0.36
	CSI	0.53	0.49	0.52	0.51	0.47	0.42	0.39	0.41	0.47	0.53	0.51
CHIRPS	POD	0.80	0.77	0.79	0.81	0.82	0.80	0.77	0.77	0.81	0.79	0.78
	FAR	0.40	0.44	0.39	0.44	0.45	0.49	0.54	0.52	0.47	0.38	0.44
	CSI	0.52	0.48	0.52	0.50	0.49	0.45	0.41	0.42	0.47	0.53	0.48
CPC	POD	0.76	0.84	0.88	0.77	0.82	0.78	0.73	0.72	0.83	0.71	0.70
	FAR	0.35	0.46	0.14	0.43	0.41	0.46	0.50	0.44	0.48	0.36	0.39
	CSI	0.54	0.49	0.77	0.49	0.52	0.47	0.42	0.46	0.47	0.51	0.49
SAOBS	POD	0.73	0.80	0.70	0.79	0.98	0.78	0.77	0.73	0.69	0.69	-
	FAR	0.36	0.36	0.33	0.37	0.18	0.41	0.53	0.42	0.48	0.42	-
	CSI	0.52	0.55	0.52	0.54	0.80	0.51	0.42	0.48	0.42	0.46	-
TRMM	POD	0.68	0.66	0.67	0.66	0.61	0.57	0.55	0.61	0.65	0.67	0.70
	FAR	0.31	0.35	0.31	0.34	0.36	0.40	0.44	0.44	0.38	0.29	0.38
	CSI	0.52	0.49	0.52	0.49	0.46	0.41	0.38	0.41	0.47	0.52	0.49

7 Discussion

This study evaluated the eight GPPs: five gauge-based products (APHRODITE, CPC, CRU, GPCC, and SAOBS), one satellite-based product (TRMM), and two reanalyzed products (CMORPH and CHIRPS) for the two sub-basins (NRB and VMD) of LMRB. Contrasting performance of GPPs were observed for these basins. The performance differences between the GPPs across SEA is also reported by Alexander et al. (2020). APHRODITE performed well for the NRB while TRMM outweighed others in VMD. The performance of APHRODITE outperforming other dataset for the NRB is consistent with the findings of Tian et al. (2021) and Tang et al. (2019). The better performance of APHRODITE in the NRB could be the outcome of the better interpolation technique addressing the topographic variation (Yi Yang et al. 2017) and was developed specifically for Monsoon Asia by taking observed gauge-based precipitation products from within the region (Yatagai et al. 2012). Saying that, the APHRODITE failed to capture the EIs in many instances for both the sub-basins. For example, the APHRODITE overestimated (underestimated) CWD (CDD). The higher CWD is the reflection of spatial average of gauge data over the grid resolution (Hussain et al. 2018; Satgé et al. 2016). In contrast, the performance of GPPs is quite similar (Sun et al. 2018) and discrepancies are not large for VMD as indicated by the statistical indices of GPPs except for SAOBS. Though the SAOBS was developed particularly for SEA with data obtained from the SEA countries, it observed the poorest skill in estimating the precipitation across VMD. This possibly could be that the gauge records might not have been made available from VMD during the production of SAOBS product.

In terms of representing the mean precipitation, gauge-based products (except GPCC for NRB and VMD and CRU only for VMD) and CHIRPS (reanalyzed product) mostly underpredicted for 75% of stations. The underestimation by these products could be attributed to the limited number of in-situ stations used for developing gridded dataset and inadequate quality control of the real-time GTS datasets whereby sometimes missing values are replaced with 0 especially for APHRODITE (Yatagai et al. 2012). The underprediction of precipitation by CPC over Myanmar and neighboring regions is also reported by Kim et al. (2019). In contrast, the GPCC overestimated (though close to observed) the mean monthly precipitation. The GPCC at the monthly scale outperforms in terms of capturing the mean monthly precipitation attributed to its increased number of stations and improved quality controlled data and the change in interpolation technique from kriging to spheremap (Alexander et al. 2020; Schneider et al. 2014). Similarly, the overestimation of precipitation of TRMM can be linked to ignorance of altitudinal variations in the algorithm, IR sensor estimating precipitation from non-raining cirrus cloud (Scheel et al. 2011) and its bias correction technique that employs GPCC as reference

dataset at monthly scale (Trinh-Tuan et al. 2019). Likewise, the higher negative bias of the CMOPRH can be linked to its inherent characteristics of underestimating the precipitation at higher altitude (Hobouchian et al. 2017) and the use of CPC (which is already underestimating precipitation) as the reference dataset for bias correction (Xie et al. 2017).

The CMORPH usually simulates light-rain events at daily scale and fewer heavy rain-events. This can be attributed to its gridding algorithm which is bilinear interpolation (Yu et al. 2009). It ranked last in terms of representing CDD and CWD for almost all the stations. It is also noteworthy to contemplate that the precipitation by the GPPs is the mean areal value of the grid boxes and the spatial resolution of the different GPPs are different which compromises the spatial representativeness of the station to the grid. This may differ significantly in terms of elevation as well. Hence, it is difficult for a station to represent the precipitation of different spatial resolution of different GPPs.

The gauge-based GPPs exhibited higher POD compared to satellite-based and reanalyzed GPPs. In contrast satellite-based and reanalyzed products were far skillful in minimizing FAR compared to the gauge-based GPPs (except for SAOBS). Similar findings were reported by Yi Yang et al. (2017). The lower POD and higher FAR for satellite and reanalyzed product can be attributed to the low ability to discriminate the drizzle and frozen precipitation during dry season (Wu et al. 2019). The higher FAR for the gauge-based product could be caused by the differences in grid location and gauge location and areal averaged value.

8 Conclusions

An evaluation and ranking of eight state-of-art GPPs were carried out for the two contrasting sub-basins of the LMRB: (i) NRB of Lao PDR with undulating terrain and (ii) humid VMD characterized by flat terrain and tropical climate. The noticeable differences in the performance of GPPs were revealed for two sub-basins. Most of former studies attempted to evaluate and recommend a best GPP for the MRB or LMRB as a whole. However, this may not be always applicable as the study domain might only be the sub-basins of smaller areal coverage. Depending upon the area considered for the study and the gauge-station availability, a diverse performance can be expected for different GPP with in sub-basins of the same basin as well. This study also indicates that there is no-size-fits-all GPP for different sub-basins for application within the LMRB as APHRODITE outperformed other GPPs at daily scale for NRB while TRMM ranked first for VMD. Further, Alexander et al. (2020) also pointed out that the differences between the performances of GPPs across the SEA is prominent. The differences in the performances can be attributed to the sparse station

density considered for developing the GPPs, station's locations, gridding algorithms, location of stations and grid point, etc.

This study employed the ECP, an automated multi-criteria decision-making approach for evaluation and ranking of the GPPs. This is also a first attempt to rank the dataset using compromise programming in these two sub-basins. This method calculates the weight based on the uncertainty with in the timeseries and rank them with reference to minimum distance (minimum error/uncertainty) from the observed dataset by limiting the interference of the researcher. Most of the studies (A. Chen et al. 2018; C.-J. Chen et al. 2017; Dandridge et al. 2019; Dang Dinh et al. 2020) that has been carried out to evaluate GPPs in the LMRB rely only on the statistical and categorical indices and the ranking is solely based on the skill of the researcher to weigh these indices. As different statistical indices may favor different GPP resulting in a conflicting choice and it is entirely dependent on skill of researcher to choose the best one. This limitation can be overcome by employing the ECP which weighs the uncertainties in statistical indices and rank GPP with no influence from the researcher. However, the performance might be affected by the number of criteria being considered and the weight obtained from entropy gets redistributed on addition or removal of evaluation indices influencing the GPP selection. Hence, care should be taken in limiting the number of statistical indices to be used. Saying that, the ECP can be a suitable choice in ranking and evaluating a large suite of GPPs and when a researcher has hard time in selecting a suitable choice based on statistical indices only.

Saying that, the ranked GPP from this study can be used for different hydro-meteorological application such as water resource planning and estimating potential for hydropower potential. The GPPs are the gateways to fill in the data gaps and spatial coverage for several hydrometeorological studies (Morales-Moraga et al. 2019; Prajapati et al. 2021). However, it should also be noted that the choice of the GPP should be considered based upon the application or purpose of the study as the error in the daily precipitation frequencies and intensities of GPPs significantly impact the output of hydrologic simulations (Luo et al. 2019). For instance, though APHRODITE was ranked best for NRB, its skill to reflect the EI for some instances (CWD, CDD) was poor. Hence, in case of extreme analysis, GPP that captures the EI well or an ensemble of the GPPs addressing the large uncertainty among the dataset shall be considered.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00704-022-04268-1>.

Acknowledgements Authors would like to extend their sincere thanks to the “Strengthening Groundwater Governance in Rapidly Urbanising Areas of the Lower Mekong Region (GIRA)” project funded by Stockholm Environment Institute (SEI) under the SUMERNET 4 All Programme which is funded by the Swedish International Development Cooperation Agency (Sida).

Author contributions SD: conceptualization, data curation, formal analysis, investigation, methodology, visualization and validation, drafting original manuscript. SS: PI, overall supervision, validation, review and editing. SK: data curation, review and editing. TPV: data curation, review and editing. ADG: review and editing. TPLN: review and editing.

Funding This work was funded by Stockholm Environment Institute (SEI) under the SUMERNET 4 All Programme.

Data availability The datasets generated during and/or analyzed during the current study are available upon reasonable request from the corresponding author.

Declarations

Conflict of interest The authors declare no competing interests.

Ethics approval All authors have read and approved this manuscript.

Consent to participate The authors have provided their consent to submit this manuscript to TAAC.

Consent for publication The authors give permission to publish this manuscript.

References

- Ahmed K, Shahid S, Wang X, Nawaz N, & Khan N. (2019). Evaluation of gridded precipitation datasets over arid regions of Pakistan. In *Water* 11(2). <https://doi.org/10.3390/w11020210>
- Alexander LV, Bador M, Roca R, Contractor S, Donat MG, Nguyen PL (2020) Intercomparison of annual precipitation indices and extremes over global land areas from in situ, space-based and reanalysis products. *Environ Res Lett* 15(5):55002. <https://doi.org/10.1088/1748-9326/ab79e2>
- Aliyar Q, Dhungana S, Shrestha S (2021) Spatio-temporal trend mapping of precipitation and its extremes across Afghanistan (1951–2010). *Theoret Appl Climatol*. <https://doi.org/10.1007/s00704-021-03851-2>
- Arshad M, Ma X, Yin J, Ullah W, Liu M, Ullah I (2021) Performance evaluation of ERA-5, JRA-55, MERRA-2, and CFS-2 reanalysis datasets, over diverse climate regions of Pakistan. *Weather Clim Extremes* 33:100373. <https://doi.org/10.1016/j.wace.2021.100373>
- Awange JL, Ferreira VG, Forootan E, Khandu, Andam-Akorful SA, Agutu NO, He XF (2016) Uncertainties in remotely sensed precipitation data over Africa. *Int J Climatol* 36(1):303–323. <https://doi.org/10.1002/joc.4346>
- Brans JP, Vincke P, Mareschal B (1986) How to select and how to rank projects: The Promethee method. *Eur J Oper Res* 24(2):228–238. [https://doi.org/10.1016/0377-2217\(86\)90044-5](https://doi.org/10.1016/0377-2217(86)90044-5)
- 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>
- Chen C-J, Senarath SUS, Dima-West IM, Marcella MP (2017) Evaluation and restructuring of gridded precipitation data over the Greater Mekong Subregion. *Int J Climatol* 37(1):180–196. <https://doi.org/10.1002/joc.4696>
- Chen A, Chen D, Azorin-Molina C (2018) Assessing reliability of precipitation data over the Mekong River Basin: A comparison of ground-based, satellite, and reanalysis datasets. *Int J Climatol* 38(11):4314–4334. <https://doi.org/10.1002/joc.5670>

- Dandridge C, Lakshmi V, Bolten J, & Srinivasan R (2019). Evaluation of satellite-based rainfall estimates in the Lower Mekong River Basin (Southeast Asia). In *Remote Sensing* 11:(22). <https://doi.org/10.3390/rs11222709>
- Dang Dinh K, Ngoc Anh T, Nguyen N, Bui D, Srinivasan R (2020) Evaluation of grid-based rainfall products and water balances over the Mekong River Basin. *Remote Sensing* 12:21–32. <https://doi.org/10.3390/rs12111858>
- Darand M, Khandu K (2020) Statistical evaluation of gridded precipitation datasets using rain gauge observations over Iran. *J Arid Environ* 178:104172. <https://doi.org/10.1016/j.jaridenv.2020.104172>
- Duan Z, Liu J, Tuo Y, Chiogna G, Disse M (2016) Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Sci Total Environ* 573:1536–1553. <https://doi.org/10.1016/j.scitotenv.2016.08.213>
- Duy NL, Nguyen TVK, Nguyen DV, Tran AT, Nguyen HT, Heidebüchel I, Merz B, Apel H (2021) Groundwater dynamics in the Vietnamese Mekong Delta: Trends, memory effects, and response times. *J Hydrol: Regional Stud* 33:100746. <https://doi.org/10.1016/j.ejrh.2020.100746>
- Ebert EE, Janowiak JE, Kidd C (2007) Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull Am Meteor Soc* 88(1):47–64. <https://doi.org/10.1175/BAMS-88-1-47>
- Fallah A, Rakhshandehroo GR, Berg POS, Orth R (2020) Evaluation of precipitation datasets against local observations in southwestern Iran. *Int J Climatol* 40(9):4102–4116. <https://doi.org/10.1002/joc.6445>
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* 2(1):150066. <https://doi.org/10.1038/sdata.2015.66>
- Gershon M, & Duckstein L (1983). Multiobjective approaches to river basin planning. *J Water Res Planning Manag-Asce - J WATER RESOUR PLAN MAN-ASCE*, 109 [https://doi.org/10.1061/\(ASCE\)0733-9496\(1983\)109:1\(13\)](https://doi.org/10.1061/(ASCE)0733-9496(1983)109:1(13))
- Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int J Climatol* 34(3):623–642. <https://doi.org/10.1002/joc.3711>
- Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* 7(1):109. <https://doi.org/10.1038/s41597-020-0453-3>
- Hobeichi S, Abramowitz G, Contractor S, Evans J (2020) Evaluating precipitation datasets using surface water and energy budget closure. *J Hydrometeorol* 21(5):989–1009. <https://doi.org/10.1175/JHM-D-19-0255.1>
- Hobouchian MP, Salio P, García Skabar Y, Vila D, Garreaud R (2017) Assessment of satellite precipitation estimates over the slopes of the subtropical Andes. *Atmos Res* 190:43–54. <https://doi.org/10.1016/j.atmosres.2017.02.006>
- Hu Z, Zhou Q, Chen X, Li J, Li Q, Chen D, Liu W, Yin G (2018) Evaluation of three global gridded precipitation data sets in central Asia based on rain gauge observations. *Int J Climatol* 38(9):3475–3493. <https://doi.org/10.1002/joc.5510>
- Huffman GJ, & Bolvin DT (2018). *26 April 2018 TRMM 3B42_3B43_doc 1*. November 2013, 1–48.
- Hussain Y, Satgé F, Hussain MB, Martinez-Carvajal H, Bonnet M-P, Cárdenas-Soto M, Roig HL, Akhter G (2018) Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theoret Appl Climatol* 131(3):1119–1132. <https://doi.org/10.1007/s00704-016-2027-z>
- Irannezhad M, Liu J (2022) Evaluation of six gauge-based gridded climate products for analyzing long-term historical precipitation patterns across the Lancang-Mekong River Basin. *Geography Sustain* 3(1):85–103. <https://doi.org/10.1016/j.geosus.2022.03.002>
- Kim I-W, Oh J, Woo S, Kripalani RH (2019) Evaluation of precipitation extremes over the Asian domain: observation and modelling studies. *Clim Dyn* 52(3):1317–1342. <https://doi.org/10.1007/s00382-018-4193-4>
- Komaragiri SR, Kumar DN (2014) Ranking of global climate models for India using multicriterion analysis. *Climate Res* 60:103–117. <https://doi.org/10.3354/cr01222>
- Le PVV, Pham HV, Bui LK, Tran AN, Pham CV, Nguyen GV, Tran PA (2021) Responses of groundwater to precipitation variability and ENSO in the Vietnamese Mekong Delta. *Hydrol Res* 52(6):1280–1293. <https://doi.org/10.2166/nh.2021.024>
- Lee GW (2006) Sources of errors in rainfall measurements by polarimetric radar: variability of drop size distributions, observational noise, and variation of relationships between R and polarimetric parameters. *J Atmos Oceanic Tech* 23(8):1005–1028. <https://doi.org/10.1175/JTECH1899.1>
- Luo X, Wu W, He D, Li Y, Ji X (2019) Hydrological simulation using TRMM and CHIRPS precipitation estimates in the Lower Lancang-Mekong River Basin. *Chin Geogra Sci* 29(1):13–25. <https://doi.org/10.1007/s11769-019-1014-6>
- Meema T, Tachikawa Y, Ichikawa Y, Yorozu K (2021) Uncertainty assessment of water resources and long-term hydropower generation using a large ensemble of future climate projections for the Nam Ngum River in the Mekong Basin. *J Hydrol: Regional Stud* 36:100856. <https://doi.org/10.1016/j.ejrh.2021.100856>
- Morales-Moraga D, Meza FJ, Miranda M, Gironás J (2019) Spatio-temporal estimation of climatic variables for gap filling and record extension using Reanalysis data. *Theoret Appl Climatol* 137(1):1089–1104. <https://doi.org/10.1007/s00704-018-2653-8>
- Opricovic S, Tzeng G-H (2004) Compromise solution by MCDM methods: a comparative analysis of VIKOR and TOPSIS. *Eur J Oper Res* 156:445–455. [https://doi.org/10.1016/S0377-2217\(03\)00020-1](https://doi.org/10.1016/S0377-2217(03)00020-1)
- Piesse M (2016). Livelihoods and food security on the Mekong River. *Future Direction International, Strategic Analysis Paper, (May 26, 2016)*. Accessed June, 13.
- Prajapati R, Talchabhadel R, Silwal P, Upadhyay S, Ertis B, Thapa B, Davids J (2021). Less rain and rainy days—lessons from 45 years of rainfall data (1971–2015) in the Kathmandu Valley, Nepal. *Theor Appl Climatol* 145 <https://doi.org/10.1007/s00704-021-03706-w>
- Prakash S, Mitra AK, Rajagopal EN, Pai DS (2016) Assessment of TRMM-based TMPA-3B42 and GSMaP precipitation products over India for the peak southwest monsoon season. *Int J Climatol* 36(4):1614–1631. <https://doi.org/10.1002/joc.4446>
- 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 MP, Seyler F, Rouché N, Pierron F, Paturol JE (2020). Evaluation of 23 gridded precipitation datasets across West Africa. *J Hydrol* 581(November 2019). <https://doi.org/10.1016/j.jhydrol.2019.124412>
- Scheel MLM, Rohrer M, Huggel C, Santos Villar D, Silvestre E, Huffman GJ (2011) Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol Earth Syst Sci* 15(8):2649–2663. <https://doi.org/10.5194/hess-15-2649-2011>
- 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. *Theoret Appl Climatol* 115(1):15–40. <https://doi.org/10.1007/s00704-013-0860-x>

- Sebesvari Z, Le HTT, Van Toan P, Arnold U, & Renaud FG (2012). *Agriculture and Water Quality in the Vietnamese Mekong Delta BT - The Mekong Delta System: Interdisciplinary Analyses of a River Delta* (F. G. Renaud & C. Kuenzer (eds.); 331–361). Springer Netherlands. https://doi.org/10.1007/978-94-007-3962-8_13
- Shen Z, Yong B, Gourley JJ, Qi W, Lu D, Liu J, Ren L, Hong Y, Zhang J (2020) Recent global performance of the Climate Hazards group Infrared Precipitation (CHIRP) with Stations (CHIRPS). *J Hydrol* 591:125284. <https://doi.org/10.1016/j.jhydrol.2020.125284>
- 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>
- Tang X, Zhang J, Wang G, Yang Q, Yang Y, Guan T, Liu C, Jin J, Liu Y, Bao Z (2019) Evaluating suitability of multiple precipitation products for the Lancang River Basin. *Chin Geogra Sci* 29(1):37–57. <https://doi.org/10.1007/s11769-019-1015-5>
- Tang X, Zhang J, Wang G, Ruben GB, Bao Z, Liu Y, Liu C, & Jin J (2021). Error Correction of Multi-Source Weighted-Ensemble Precipitation (MSWEP) over the Lancang-Mekong River Basin. In *Remote Sensing* 13: (2). <https://doi.org/10.3390/rs13020312>
- Tarek M, Brissette F, Arsenault R (2021) Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. *Hydrol Earth Syst Sci* 25(6):3331–3350. <https://doi.org/10.5194/hess-25-3331-2021>
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res: Atmospheres* 106(D7):7183–7192. <https://doi.org/10.1029/2000JD900719>
- Tian W, Liu X, Wang K, Bai P, Liang K, Liu C (2021) Evaluation of six precipitation products in the Mekong River Basin. *Atmos Res* 255:105539. <https://doi.org/10.1016/j.atmosres.2021.105539>
- Trinh-Tuan L, Matsumoto J, Ngo-Duc T, Nodzu MI, Inoue T (2019) Evaluation of satellite precipitation products over Central Vietnam. *Prog Earth Planet Sci* 6(1):54. <https://doi.org/10.1186/s40645-019-0297-7>
- Try S, Tanaka S, Tanaka K, Sayama T, Oeurng C, Uk S, Takara K, Hu M, Han D (2020) Comparison of gridded precipitation datasets for rainfall-runoff and inundation modeling in the Mekong River Basin. *PLoS ONE* 15(1):e0226814–e0226814. <https://doi.org/10.1371/journal.pone.0226814>
- Upadhyay S, Silwal P, Prajapati R, Talchabhadel R, Shrestha S, Duwal S, Lakhe H (2022). Evaluating magnitude agreement and occurrence consistency of CHIRPS product with ground-based observations over medium-sized river basins in Nepal. In *Hydrology* 9:(8). <https://doi.org/10.3390/hydrology9080146>
- van den Besselaar EJM, van der Schrier G, Cornes RC, Iqbal AS, Klein Tank AMG (2017) SA-OBS: A Daily Gridded Surface Temperature and Precipitation Dataset for Southeast Asia. *J Clim* 30(14):5151–5165. <https://doi.org/10.1175/JCLI-D-16-0575.1>
- Wijngaard JB, Klein Tank AMG, Können GP (2003) Homogeneity of 20th century European daily temperature and precipitation series. *Int J Climatol* 23(6):679–692. <https://doi.org/10.1002/joc.906>
- Wu W, Li Y, Luo X, Zhang Y, Ji X, Li X (2019) Performance evaluation of the CHIRPS precipitation dataset and its utility in drought monitoring over Yunnan Province, China. *Geomat Nat Haz Risk* 10(1):2145–2162. <https://doi.org/10.1080/19475705.2019.1683082>
- Xie P, Chen M, Yang S, Yatagai A, Hayasaka T, Fukushima Y, Liu C (2007) A Gauge-Based Analysis of Daily Precipitation over East Asia. *J Hydrometeorol* 8(3):607–626. <https://doi.org/10.1175/JHM583.1>
- 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(6):1617–1641. <https://doi.org/10.1175/JHM-D-16-0168.1>
- Yang Y, Wang G, Wang L, Yu J, Xu Z (2014) Evaluation of gridded precipitation data for driving SWAT model in area upstream of Three Gorges Reservoir. *PLoS ONE* 9(11):e112725
- Yang Yi, Tang J, Xiong Z, Dong X (2017) Evaluation of high-resolution gridded precipitation data in arid and semiarid regions. *J Hydrometeorol* 18(12):3075–3101
- Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A (2012) APHRODITE: constructing a long-term daily gridded precipitation dataset for asia based on a dense network of rain gauges. *Bull Am Meteor Soc* 93(9):1401–1415. <https://doi.org/10.1175/BAMS-D-11-00122.1>
- Yong B, Liu D, Gourley JJ, Tian Y, Huffman GJ, Ren L, Hong Y (2015) Global view of real-time TRMM multisatellite precipitation analysis: implications for its successor global precipitation measurement mission. *Bull Am Meteor Soc* 96(2):283–296. <https://doi.org/10.1175/BAMS-D-14-00017.1>
- Yu Z, Yu H, Chen P, Qian C, Yue C (2009) Verification of tropical cyclone-related satellite precipitation estimates in mainland China. *J Appl Meteorol Climatol* 48(11):2227–2241. <https://doi.org/10.1175/2009JAMC2143.1>
- Zandler H, Haag I, Samimi C (2019) Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. *Sci Rep* 9(1):15118. <https://doi.org/10.1038/s41598-019-51666-z>
- Zeleny M (1973). “Compromise programming” Multiple criteria decision making. *Compromise Programming in Multiple Criteria Decision Making*, 262–301.
- Zhang X, Alexander L, Hegerl GC, Jones P, Tank AK, Peterson TC, Trewin B, Zwiers FW (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wires Clim Change* 2(6):851–870. <https://doi.org/10.1002/wcc.147>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.