

## Research Article

# Evaluating the Performance of Secondary Precipitation Products through Statistical and Hydrological Modeling in a Mountainous Tropical Basin of India

Kolluru Venkatesh <sup>1</sup>, N. Y. Krakauer <sup>2</sup>, E. Sharifi <sup>3</sup>, and H. Ramesh <sup>1</sup>

<sup>1</sup>Department of Applied Mechanics & Hydraulics, NIT Surathkal, Mangalore, Karnataka, India

<sup>2</sup>Department of Civil Engineering and NOAA-CREST, The City College of New York, New York 10031, NY, USA

<sup>3</sup>German Research Centre for Geosciences, Potsdam, Germany

Correspondence should be addressed to Kolluru Venkatesh; [venkateshkolluru95@gmail.com](mailto:venkateshkolluru95@gmail.com)

Received 10 August 2020; Revised 5 October 2020; Accepted 29 October 2020; Published 17 November 2020

Academic Editor: Filomena Romano

Copyright © 2020 Kolluru Venkatesh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the performance of gridded rainfall datasets for precipitation detection and streamflow simulations in India's Tungabhadra river basin. Sixteen precipitation datasets categorized under gauge-based, satellite-only, reanalysis, and gauge-adjusted datasets were compared statistically against the gridded Indian Meteorological Dataset (IMD) employing two categorical and three continuous statistical metrics. Further, the precipitation datasets' performance in simulating streamflow was assessed by using the Soil and Water Assessment Tool (SWAT) hydrological model. Based on the statistical metrics, Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE) furnished very good results in terms of detecting rainfall, followed by Climate Hazards Group Infrared Precipitation (CHIRP), National Centres for Environmental Prediction-Climate Forecast System Reanalysis (NCEP CFSR), Tropical Rainfall Measurement Mission (TRMM) 3B42 v7, Global Satellite Mapping of Precipitation Gauge Reanalysis v6 (GSMaP\_Gauge\_RNL), and Multisource Weighted Ensemble Precipitation (MSWEP) datasets which had good-to-moderate performances at a monthly time step. From the hydrological simulations, TRMM 3B42 v7, CHIRP, CHIRPS 0.05°, and GSMaP\_Gauge\_RNL v6 produced very good results with a high degree of correlation to observed streamflow, while Soil Moisture 2 Rain-Climate Change Initiative (SM2RAIN-CCI) dataset exhibited poor performance. From the extreme flow event analysis, it was observed that CHIRP, TRMM 3B42 v7, Global Precipitation Climatology Centre v7 (GPCC), and APHRODITE datasets captured more peak flow events and hence can be further implemented for extreme event analysis. Overall, we found that TRMM 3B42 v7, CHIRP, and CHIRPS 0.05° datasets performed better than other datasets and can be used for hydrological modeling and climate change studies in similar topographic and climatic watersheds in India.

## 1. Introduction

Precipitation is an intrinsic component of the hydrological cycle. Whether measured directly through rain gauge stations or measured from different satellite sensors, it plays a crucial role in water resources management, climatic research, and disaster management studies. Though in situ ground-based precipitation datasets provide highly accurate results, the unavailability of data and sparse and uneven distribution of gauges over unpopulated areas makes it

challenging to use them for global applications. Recently, in [1], it was highlighted that the number of rain gauges that measure precipitation globally is surprisingly small, covering less than half a soccer field. This finite gauge station availability motivated researchers to develop and test satellite/secondary rainfall estimates available from different sources/sensors having high temporal and spatial resolutions providing global coverage at subdaily, daily, and monthly time steps. These secondary datasets have high potential in monitoring precipitation, making it compelling to use them

in water resource management studies where in situ observations are scarce, particularly in remote areas [2]. Advancement in blending infrared and microwave datasets and the availability of near-global coverage with multitemporal resolutions have increased the applicability of satellite rainfall datasets over a wide range of applications. However, a direct application of these precipitation products in hydrological modeling might lead to erroneous outputs, especially for extreme flow simulations [3]. Therefore, a comprehensive validation in replicating both the climate and hydrological components is essential to identify the best precipitation product over a specific region.

Several studies were conducted by comparing different satellite precipitation products (SPPs) with gauge-based or radar-based datasets in terms of statistical metrics evaluation or hydrological modeling to predict SPPs ability and efficiency in detecting rainfall accurately. Numerous studies related to the assessment and evaluation of SPPs against gauge-based data can be found worldwide [4–8]. Reviews related to the SPPs evaluation through various hydrological framework models can also be found in different climatic and geographical regions [9–12]. Most of these studies focused on either analyzing a single rainfall product's performance in hydrological modeling or evaluating the efficiency of a few rainfall products in runoff simulations, thus restricting their analysis to specific products [7, 13–20]. Most of the studies have not considered reanalysis products during their evaluation or have not recalibrated each rainfall dataset, thus missing to differentiate the in situ corrected and uncorrected dataset efficiencies [15, 21–24]. Studies that executed both statistical and hydrological comparisons also revealed that the precipitation datasets that prove effective in statistical comparison might not exhibit the same accuracy while performing hydrological simulations [9–11]. This mandates the assessment of SPPs by employing different hydrological models/SPPs and by conducting both hydrological and statistical analyses in a basin to truly interpret the behavior, characteristics, efficiency, and performance of an SPP for hydrological applications. Very few studies tested the efficiency of more precipitation datasets; however, these studies evaluated the SPPs performance using lumped hydrological models, thus not considering the intrinsic spatial behavior of river basin characteristics that are averaged over the subbasins [17, 25, 26]. Moreover, the few studies that implemented semidistributed/distributed models have not accommodated and tested the efficiency of multiple/more number of SPPs [18, 27–30]. This provoked us to test numerous SPPs efficiency using a semidistributed hydrological model in the current study.

Further, the SPPs produced from satellite sensors are increasingly considered for hydrological modeling because of their long-term, consistent, and continuous data availability in mountainous and hilly regions [31–33]. However, a recent study conducted in [20] in different climatic regions of India concluded that the precipitation datasets failed in detecting rainfall in the tropical region. This mandates the testing of SPPs performance by employing statistical and hydrological techniques in a mountainous tropical river basin. To overcome these shortcomings/research gaps, we

evaluated 16 daily precipitation datasets against gauge-based gridded data and employed a semidistributed hydrological model in a tropical river basin of India. The 16 precipitation datasets considered in the current study are grouped under four categories, that is, gauge-based, satellite-only, reanalysis, and gauge-adjusted datasets, according to the classification mentioned by [34, 35]. To address these research gaps, the gridded rainfall products were analyzed for the period of 2000–2012, and different statistical coefficients were computed (i) to check the performance of these datasets in detecting rainfall when compared with gauge-derived gridded data, (ii) for comparing the simulated streamflow against observed streamflow using the Soil and Water Assessment Tool (SWAT), a semidistributed hydrological model, in a hilly tropical catchment of India, and (iii) for assessing the extreme flows simulated by the SWAT model employing these SPPs. Though these datasets' performance varies across space and time, the study provides insight to researchers into selecting precipitation datasets when working on similar climatic and topographic regions for hydrological modeling and other precipitation-related studies.

The water input, that is, precipitation, was replaced with 16 different precipitation estimate products obtained from various sources to find their effects on runoff simulation using the SWAT model. Hydrological models are sensitive to input variable changes, where a small change in the input data can result in more significant deviations in output. Hence, it may be assumed that weather data that can simulate streamflow against observed streamflow with the highest correlation and least variance and bias have the best claim to be accurate. Since studies using satellite precipitation datasets on the hydrological model were scarce on mountainous tropical river basins, particularly in India, we were motivated to assess the SPPs efficiency for simulating runoff in an Indian river basin. SWAT model was selected for modeling the Tungabhadra river basin, south India, and different standard diagnostics like  $R^2$  (coefficient of determination), N-S (Nash-Sutcliffe coefficient), and PBias (percentage bias) were computed to evaluate the performance of each SPP. Though many distributed/semidistributed models such as MIKE SHE [36], TOP NET [37], PIHM [38], CREST [39], and VIC [40] were developed to incorporate the heterogeneity of basins, the physically based semidistributed model, that is, SWAT, was implemented by many researchers around the world due to its applicability and efficiency in simulating the characteristics of a basin effectively. SWAT model has been applied successfully for a wide variety of studies related to uncertainty analysis [7, 41, 42], climate change [43, 44], land-use change [45–47], hydrological modeling [48, 49], best management practices generation [50, 51], water quality analysis [52, 53], and hydropower assessment [54, 55]. A study by [56] concluded that, as of 2019, around 4370 articles were published which implemented the SWAT model in their studies for a wide range of applications in diverse catchments having varied topographic and climatic conditions. As the current study area is agriculturally dominated, SWAT can fairly simulate the results close to the observed values as SWAT is mainly

developed to model agricultural watersheds [57]. Hence, we adopted the SWAT model instead of opting for other distributed models.

## 2. Study Area

Tungabhadra, a major tributary of river Krishna, is a transboundary river shared by Andhra Pradesh and Karnataka states in south-western India. The basin has an area of around 69,552 km<sup>2</sup> with elevation ranging from 246 m to 1921 m above mean sea level [17, 45]. Tungabhadra river originates from the confluence of twin rivers Tunga and Bhadra in the Western Ghats of Karnataka at an altitude of about 610 m above mean sea level. The current study area is a part of the upper Tungabhadra river basin with a catchment area of 7778 km<sup>2</sup>, considered up to the Honnali gauge station (having elevation of 557 m), which is the outlet of the catchment, as shown in Figure 1. The current study area lies between 74°00' to 76°30' E and 13°00' to 15°30' N, as shown in Figure 1.

The upper part of the catchment has an undulating terrain of Western Ghats and receives high rainfall compared to the lower portion of the watershed. The basin receives an average rainfall of about 1024 mm per year with mean maximum temperatures and minimum temperatures ranging from 26.3°C to 35.5°C and from 13.8°C to 22.3°C, respectively, and the relative humidity varying from 17% to 92% [58].

## 3. Input Datasets for the SWAT Model

Shuttle Radar Topographic Mission (SRTM) Digital Elevation Model (DEM) was used in this study with a spatial resolution of 30 m, downloaded from USGS Earth Explorer. The slope map was generated from DEM by selecting multiple classes based on the steepness of the surface. The slope map is categorized into five levels (0–10, 10–20, 20–30, 30–40, and >40 m). The slope map was classified by the Jenks classification (natural break) scheme available in the SWAT interface. The Jenks classification clearly sorts out ranges that can qualitatively represent the study area's slope distribution. The SWAT interface allows us to define a maximum of five different classes to delineate discrete Hydrologic Response Units (HRUs). The slope ranges were designed to fit the homogenous catchment area within each slope range, looking for the most similar slope influence distribution in detecting HRUs. Soil information was obtained from the Food and Agricultural Organization (FAO) Digital Soil Map of the World (DSMW) based on the Harmonized World Soil Database (HWSD) and produced at a scale of 1 : 5,000,000 based on existing regional and national soil information. Land use land cover (LULC) was generated from Landsat-8 satellite imagery by employing maximum likelihood algorithm in ERDAS Imagine software. The LULC was classified into six classes, namely, agricultural land, cultivated land, barren land, forest area, water body, and built-up area, as shown in Supplementary Figure 1. Based on ground validation, the classified images' overall accuracy was 85.94%, with a kappa coefficient of 0.789. SWAT requires five meteorological input variables. Of these, precipitation was

taken from any of the 16 different precipitation datasets considered and temperature from IMD (Indian Meteorological Department) gridded data, whereas relative humidity, solar radiation, and wind speed were generated using the SWAT weather generator. The observed daily streamflow data were obtained from India-WRIS (Water Resources Information System) for the Honnali gauge station for 13 years spanning from 2000 to 2012. The difference in the collection time of LULC (2018) and weather parameters (2000 to 2012) is acceptable, since previous research conducted by [47] in the Tungabhadra river basin has concluded that the effect of change in LULC on streamflow simulations is negligible. Hence, for accurate classification and representation of land use classes, 2018 LULC was considered in the current study.

*3.1. Overview of Precipitation Datasets.* Different spatial resolutions for the same dataset are considered in this study to understand the effect of these variant resolutions of the models performance. A brief description of the precipitation datasets grouped under four categories, that is, gauge-based, satellite-only product, reanalysis, and gauge-adjusted datasets, is given in Table 1. More information regarding the development of SPPs can be found in the supplementary file under the section Supplementary Materials.

## 4. Methodology

Testing the capability of these datasets is generally performed in two approaches: (i) by keeping all topographic and climatic variables and sensitive parameters constant with varying precipitation datasets (traditional method) and (ii) by calibrating the model with constant topographic and climatic variables but with varying parameters for each precipitation dataset. The former approach uses a single set of parameters obtained from calibrating a standard precipitation dataset (generally station data) and implementing those sensitive parameters for calibrating and validating the model with other precipitation datasets. The latter approach deals with calibrating and validating the model separately for each input forcing (precipitation datasets).

The traditional method may induce some uncertainty, since all the datasets may not be sensitive to similar parameters. To overcome this limitation, the present study adopted the alternative methodology of calibrating the SWAT model separately with each precipitation dataset. Calibrating the model individually for each input dataset will help to drive the model to its maximum efficacy with the most suitable parameters for each input dataset. Further, the results can be validated using consistent statistical metrics. The alternate approach helps us understand the most capable precipitation dataset for simulating streamflow that accurately matches the observed streamflow.

A similar methodology was implemented by [69, 70], where each precipitation dataset is calibrated separately to assess the efficiency of precipitation datasets. Their studies proved that using the MSWEP precipitation dataset's sensitive parameters gave better results than using a standard gauge dataset's sensitive parameters (while calibrating

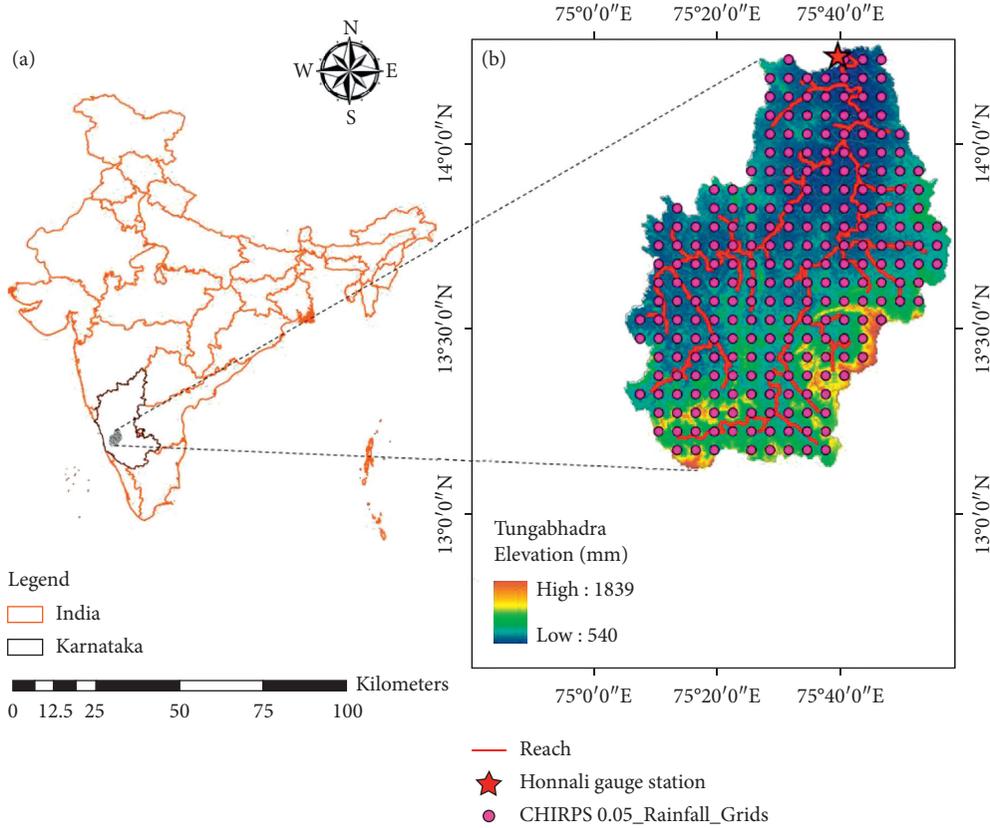


FIGURE 1: (a) Study area map of India and Karnataka. (b) DEM of Tungabhadra along with CHIRPS 0.05° rainfall data grids, Honnali gauge station, and Tungabhadra river reach.

MSWEP product). Hence, in the present study, each precipitation dataset is calibrated separately to test the dataset's maximum efficiency in producing streamflow using the SWAT model that can match against observed streamflow.

**4.1. Continuous and Categorical Statistical Indices.** The categorical statistics include probability of detection (POD) and false alarm ratio (FAR) metrics, whereas continuous statistical indices encompass correlation coefficient (CC), and root mean square error (RMSE), and bias. The precipitation detection ability can be assessed using categorical metrics, whereas SPPs' performance in estimating precipitation is determined using continuous statistical metrics. POD constitutes the ratio of hits (accurate detection of rainfall as reference data) to the actual number of rainfall events according to the base dataset (IMD). FAR represents the ratio of misses (when SPP records the rainfall during the absence of precipitation in the base dataset) to the events not diagnosed by reference dataset. The categorical metrics assessment is essential in real-time flood monitoring studies because of the importance of accurately detecting extreme precipitation. These statistics will help us to comprehend the hydrological consequences of sources of errors in SPP [71–73]. For computing POD and FAR, a threshold of 1 mm/day was implemented in the study as mentioned in [7, 20, 74, 75]. The categorical metrics were computed either for the entire time series or after segregating for different rainfall regimes. Table 2 exhibits the criterion implemented

to partition the precipitation time series into multiple components such as low, medium, and high rainfall.

CC represents the degree of synchronicity between SPP and gauge or gridded data. RMSE indexes the data accuracy or the averaged error magnitude between the gauge and SPP. Bias signifies the degree of underestimation or overestimation of simulated data concerning observed data. Metrics with lower bias and RMSE and higher CC represent higher SPP accuracy with respect to base or reference datasets. More details regarding the metrics and formulas can be found in [74, 76]. The equations employed for calculating POD, FAR, CC, RMSE, and bias are represented as follows:

$$\begin{aligned}
 \text{POD} &= \frac{H}{H + M}, \\
 \text{FAR} &= \frac{F}{H + F}, \\
 r &= \frac{\sum(G - \bar{G})(S - \bar{S})}{\sqrt{\sum(G - \bar{G})^2} \sqrt{\sum(S - \bar{S})^2}}, \\
 \text{RMSE} &= \sqrt{\frac{\sum(G - S)^2}{n}}, \\
 \text{bias} &= \frac{\sum S}{\sum G}.
 \end{aligned} \tag{1}$$

TABLE 1: Overview of the 16 daily (or subdaily) gridded precipitation datasets evaluated in this study.

S. no.	Precipitation dataset	Spatial resolution (degrees)	Spatial coverage	Temporal coverage	References
Gauge-based					
1	IMD (Indian Meteorological Department)	0.25	India	1980–2014	Pai et al. [59]
2	APHRODITE (Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation)	0.25	Monsoon Asia 60°E-150°E and 15°S-55°N	1951–2015	Yatagai et al. [60]
Satellite-only					
3	SM2RAIN-CCI (Soil Moisture 2 Rain-Climate Change Initiative)	0.25	Global	1998–2015	Ciabatta et al. [61]
Reanalysis products					
4	CHIRP v2.0 (Climate Hazards Group Infrared Precipitation version 2.0)	0.05	50°S-50°N	1981-NRT	Funk et al. [62]
5	CHIRPS v2.0 (Climate Hazards Group Infrared Precipitation with Stations version 2.0)	0.05	50°S-50°N	1981-NRT	Funk et al. [62]
6	CHIRPS v2.0 (Climate Hazards Group Infrared Precipitation with Stations version 2.0)	0.25	50°S-50°N	1981-NRT	Funk et al. [62]
7	GSMAP Gauge RNL v6 (Global Satellite Mapping of precipitation gauge reanalysis version 6)	0.1	Global	2000–2014	Okamoto et al. [63]
8	NCEP-CFSR (National Centres for Environmental Prediction-Climate Forecast System Reanalysis)	0.31	Global	1979–2014	Saha et al. [64]
9	PGF v2 (Princeton Global Forcing version 2)	0.25	90°N-90°S	1901–2012	Sheffield et al. [65]
10	PGF v2 (Princeton Global Forcing version 2)	0.5	90°N-90°S	1901–2012	Sheffield et al. [65]
11	MSWEP v1.2 (Multisource Weighted Ensemble Precipitation version 1.2)	0.25	Global	1979–2015	Beck et al. [66]
Gauge-adjusted					
12	GPCP-CDR v1.3 (Global Precipitation Climatology Project-Climate Data Record)	1	Global	1996-NRT	Huffman et al. [67]
13	GPCC v.2018 (Global Precipitation Climatology Centre version 2018)	1	Global	1982–2016	Schneider et al. [68]
14	GPCC v.7 (Global Precipitation Climatology Centre version.7)	0.5	90°N-90°S	1901–2013	Schneider et al. [65]
15	PERSIANN-CDR (Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks-Climate Data Record)	0.25	60°N-60°S	1983–2017	Schneider et al. [65]
16	TRMM 3B42 v7 (Tropical Rainfall Measurement Mission 3B42 version 7)	0.25	50°N-50°S	1998-NRT	Huffman et al. [67]

Abbreviations in the data source category column are defined as follows: G, gauge; S, satellite; and R, reanalysis. The term “global” indicates fully global coverage including ocean areas.

TABLE 2: Partition of overall precipitation time series into low, medium, and high precipitation values (PCP is precipitation,  $\mu$  is mean of precipitation, and  $\sigma$  is the standard deviation of precipitation).

Rainfall regime	Criterion
Low	$PCP < \mu$
Medium	$PCP \geq \mu$ and $PCP \leq \mu + 2\sigma$
High	$PCP > \mu + 2\sigma$

Here  $H, M, F$  represent the numbers of hits, misses, and false events of rainfall,  $G$  and  $S$  represent observed (gauge) and satellite precipitation estimates,  $\bar{G}$  and  $\bar{S}$  are averages of observed and satellite precipitation estimates, and  $n$  represents the number of data pairs.

**4.2. SWAT Model.** In this study, the SWAT model was employed for modeling the Tungbhadra river basin, India, to assess how the change in different input datasets (rainfall) affects the hydrologic process, that is, runoff. United States Department of Agriculture-Agricultural Research Service (USDA-ARS) developed a deterministic model, that is, SWAT [77], which uses both lumped (rainfall per subbasin) and distributed (HRU combination of unique soil, slope, and land use characteristics) variables for hydrological modeling. The watershed delineated using SWAT model produced 73 subbasins and 503 HRUs in the current study. SWAT uses readily available inputs for predicting various components related to water, sediment, and agricultural chemical yields for all types of watersheds at daily, monthly, and annual time step. The overall methodology is represented in Figure 2.

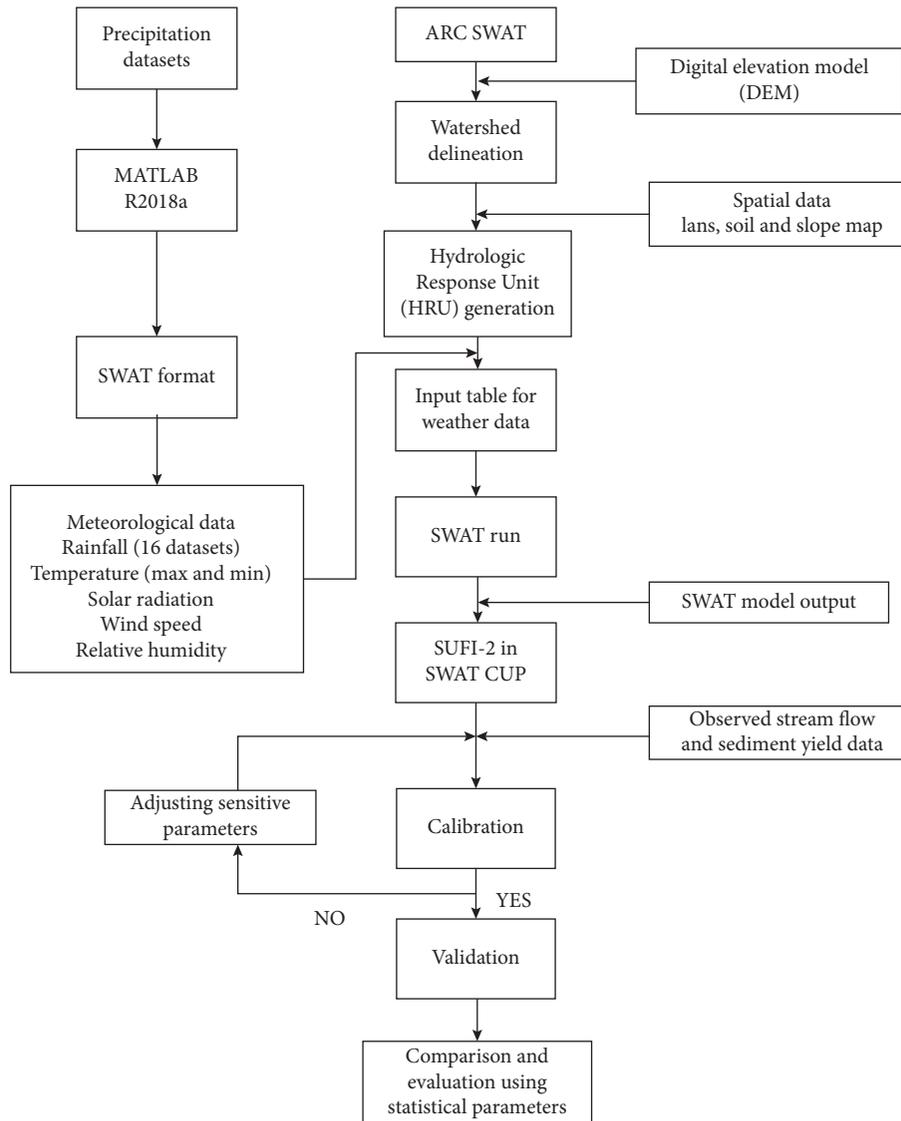


FIGURE 2: Flowchart of simulating streamflow using the SWAT model.

The SWAT model was calibrated at a monthly time step rather than being employed at a daily scale because of the catchment size and the type of land use in the basin that influences the hydrologic response times due to changes in the characteristic velocity. Urban and agriculturally dominated catchments exhibit faster and intermediate hydrological responses, whereas forest dominated basins will exhibit slower hydrological responses. The watershed in the current study is forest-dominated followed by agriculture (part of agriculture has similar characteristics of forest due to higher occupancy of areca nut and coconut plantation in the study area), which has slower-to-intermediate hydrologic response time. Hence, calibrating a model with different inputs may necessitate the observations whose temporal resolutions correlate with respective hydrologic response times implying that, for significant results, the model should be calibrated at a monthly time step [78]. Many studies [42, 74, 79–81] conducted over different regions and

catchments and two studies [47, 58] over the Tungabhadra basin even concluded that the SWAT model's performance in simulating the flows is excellent at a monthly time step when compared to a daily time scale. Hence, the current study employed the SWAT model to simulate streamflows with different precipitation datasets at a monthly time step.

**4.3. SWAT CUP.** Calibration and validation of SWAT output (runoff) were carried out using the SUFI-2 (Sequential Uncertainty Fitting) algorithm in SWAT CUP. SWAT CUP is an autocalibration and uncertainty analysis tool that optimizes a range of input parameters iteratively through calibration batches. The autocalibration that depends on the global optimization method, that is, Shuffled Complex Evolution developed at the University of Arizona (SCE-UA), is implemented in the current study to optimize the sensitive parameters. SWAT CUP provides various

sensitive parameters with varying minimum and maximum values, which should be adjusted to obtain the best fit values when compared with observed data.

As large numbers of parameters can potentially be set within the SWAT model, sensitive parameters are identified based on previous literature for performing sensitivity analysis. The SWAT model was run for 1000 simulations with this list of candidate sensitive parameters. We selected the sensitive parameters, according to  $t$ -statistic and  $p$  value. The  $p$  value ranges from 0 to 1, which provides the significance of sensitivity, where a value closer to 1 is identified as a nonsensitive parameter and vice versa. The sensitive parameters (listed in Supplementary Table 1) obtained after sensitivity analysis were employed to calibrate and validate the model separately with each input forcing. The calibration of the SWAT model for the study area was carried out by comparing observed monthly streamflows over the training period with simulated values at the outlet of the basin by maintaining Nash-Sutcliffe (N-S) coefficient as the objective function. Around 14 parameters with varying ranges are employed for a different set of iterations for different input dataset calibrations. The various parameters adjusted in the present study are CN2 (SCS runoff curve number), Alpha\_BF (base flow alpha-factor), GW\_Delay (the delay in groundwater to resurface), GWQMN (depth of base flow alpha factor for bank storage), CH\_N2 (Manning “ $n$ ” coefficient for main channel), CH\_K2 (main channel’s hydraulic conductivity), SOL\_AWC (water capacity available for the soil layer), SOL\_K (saturated hydraulic conductivity of soil), ESCO (soil evaporation compensation factor), GW\_REVAP (groundwater “revap” coefficient), REVAPMN (depth of water required for “revap” to occur in shallow aquifer), SLSUBBSN (average slope length), SLSOIL (slope length for lateral subsurface flow), and ALPHA\_BNK (base flow alpha factor for bank storage). The allowable ranges and obtained fitted values are provided in Supplementary Table 1. The model was calibrated for the period of 2002–2008 and validated for the period of 2009–2012 with a warm-up period of 2 years. The warm-up period is required to train the model according to the datasets and attain hydrological parameter balance from the initial state.

## 5. Results and Discussion

### 5.1. Spatial Patterns of Sixteen Precipitation Datasets.

Figure 3 depicts the spatial variations of the 13-year mean annual precipitation of the sixteen precipitation datasets over the Tungabhadra river basin. All the precipitation datasets exhibited an increasing rainfall trend from northeast to southwest over the Tungabhadra river basin. IMD, which was considered as standard, has a mean annual rainfall varying across the basin from 795 mm to 3195 mm. Another gauge dataset, APHRODITE (Figure 3(a)), estimated an average areal rainfall almost similar to IMD ranging from 667 mm to 2917 mm.

The mean areal precipitations depicted by CHIRP (Figure 3(b)) and CHIRPS 0.25° (Figure 3(d)) under the reanalysis category were close to the ranges of IMD (Figure 3(i)). In comparison, CHIRPS 0.05° (Figure 3(c)) and

NCEP-CFSR (Figure 3(k)) had higher precipitation values compared to IMD. Reanalysis products PGF 0.5°, PGF 0.25°, and MSWEP (Figures 3(m), 3(n) and 3(j)) projected very low rainfall magnitudes over the study area compared to IMD.

GPCC v2018 and GPCC v7 (Figures 3(e) and 3(f)) have higher values of mean areal precipitation when compared to IMD, while lower average annual values were seen for TRMM 3B42 v7 (Figure 3(p)). GPCP-CDR v1.3 (Figure 3(g)) and PERSIANN-CDR (Figure 3(l)) datasets under the gauge-adjusted category were unable to detect the rainfall variation over the study area with precipitation spatially varying from 867 to 1036 mm and from 698 to 917 mm, respectively. The substantial difference of rainfall amount between southwest and northeast parts of the basin signifies the orographic effect of precipitation due to the existence of the Western Ghats mountains in the upper part of the basin, that is, at the southwest side of the catchment. Overall, GPCP-CDR v1.3, MSWEP, PERSIANN-CDR, PGF 0.25°, and SM2RAIN-CCI datasets failed to represent patterns over the study area accurately. CHIRP, CHIRPS 0.05°, and NCEP-CFSR captured the patterns effectively when compared against the IMD precipitation dataset.

### 5.2. Evaluation of Satellite Precipitation Datasets through Statistical Measures

#### 5.2.1. Continuous Statistical Indices

(1) *From Gauge-Based and Satellite-Only Datasets.* The correlation coefficient for the 15 datasets compared with the IMD mean precipitation distribution is close to their corresponding perfect values (above 0.75), except for PERSIANN-CDR, which has a weak correlation (Figure 4). Under the gauge-based dataset category, APHRODITE has lower RMSE (2.89) and moderate overestimated bias (0.99). The satellite-derived SM2RAIN-CCI dataset provided a good correlation against the IMD dataset at a monthly time step (Figure 4). In contrast, the other two statistical metrics exhibited significant overestimation and higher RMSE values for the SM2RAIN-CCI dataset.

(2) *From Reanalysis Datasets.* All datasets under the reanalysis category exhibited a high correlation against the IMD dataset at the monthly time step (Figure 4). The underestimation of precipitation amount by the CHPClim (CHIRP, CHIRPS 0.25°, and CHIRPS 0.05°) datasets has been reported earlier by [42, 73, 82]. The NCEP-CFSR dataset under the reanalysis category revealed better performance with high CC (0.86), low RMSE (3.33), and low bias (−0.77) (Table 3). The Princeton datasets (PGF 0.25° and PGF 0.5°) overestimated the precipitation magnitudes, with moderate bias (1.93 and 1.46) and moderate RMSE (4.39 and 3.69 mm).

(3) *From Gauge-Adjusted Datasets.* TRMM 3B42 v7 and GPCC v7 demonstrated effective performance in detecting precipitation similar to IMD as they exhibited near-perfect results in terms of CC, RMSE, and bias under gauge-adjusted category at a monthly time step (Table 3). The close

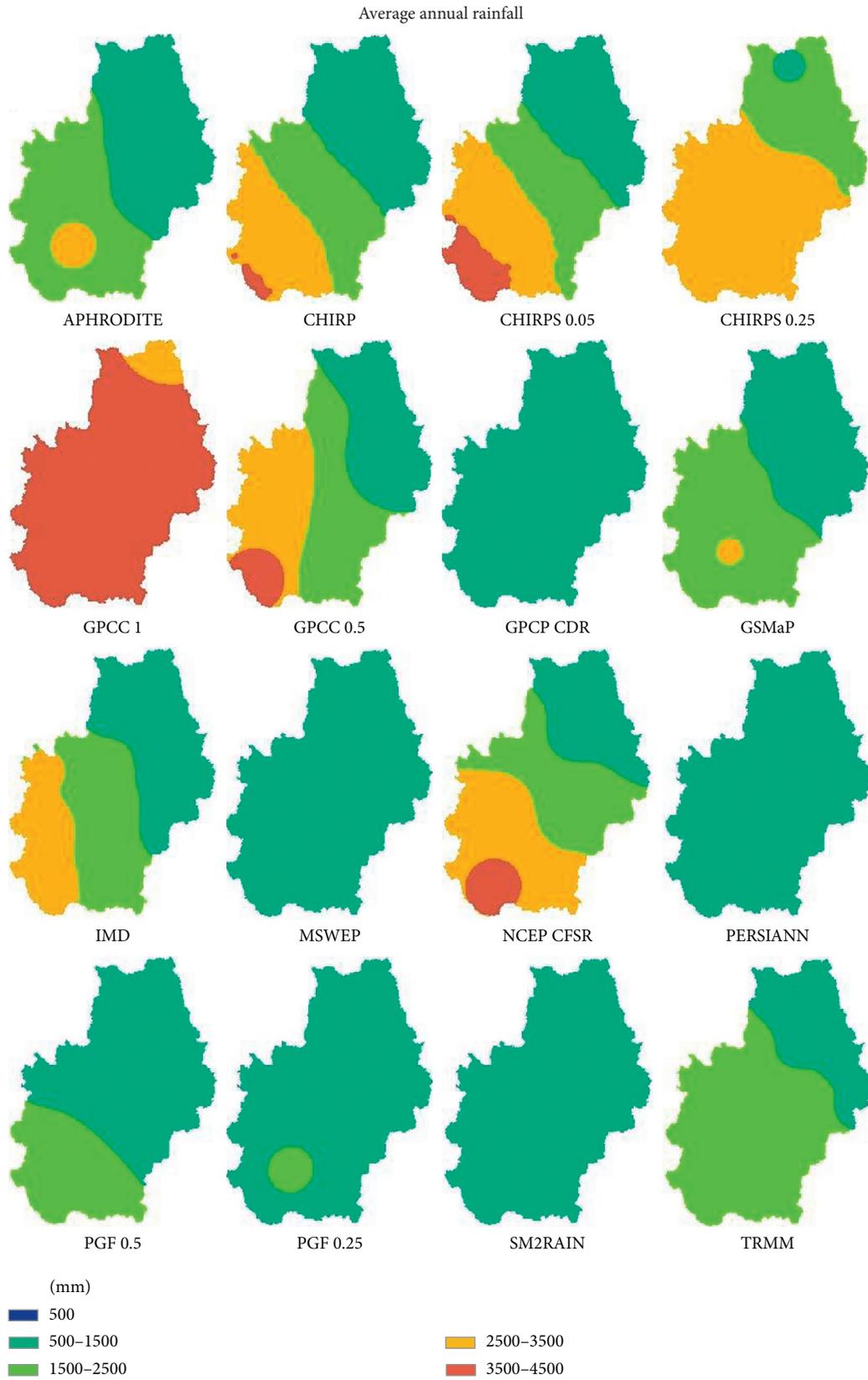


FIGURE 3: Spatial variations of the thirteen-year mean annual precipitation of sixteen rainfall datasets.

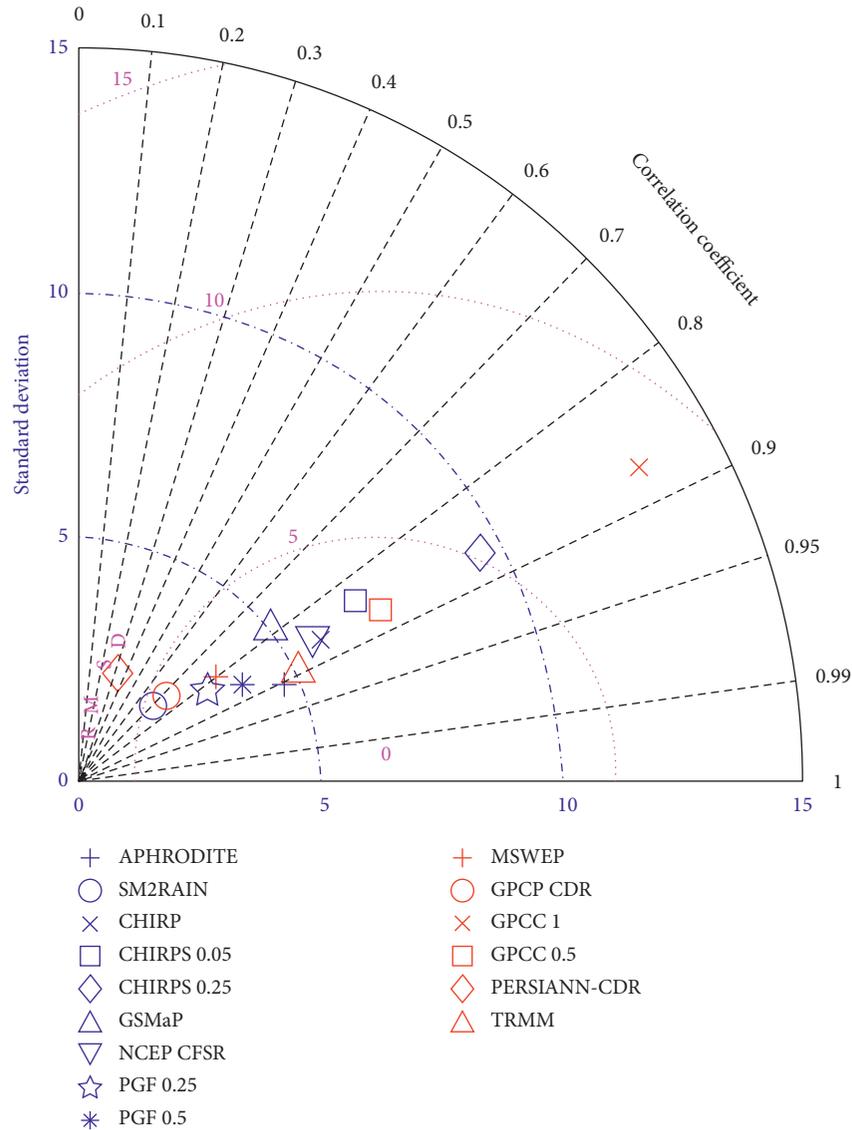


FIGURE 4: Taylor diagram depicted for all 15 gridded rainfall datasets against IMD gridded dataset at monthly time step.

agreement and low bias between the TRMM 3B42 v7 and IMD datasets can be attributed to the gauge-based dataset’s pivotal role in the bias removal process [83]. Poor performance and overestimation were observed from the statistical values of PERSIANN-CDR with high error magnitudes (RMSE = 6.42) and weak correlation (0.34).

Overall, APHRODITE furnished excellent results in detecting rainfall with less bias and error magnitudes than the IMD dataset from the results of continuous statistical metrics. CHIRP, NCEP-CFSR, TRMM 3B42 v7, GSMaP\_Gauge\_RNL v6, and MSWEP datasets exhibited good-to-moderate performances at a monthly time step.

### 5.2.2. Categorical Statistical Metrics

(1) *From Gauge-Based and Satellite-Only Datasets.* Categorical statistics, that is, POD and FAR, are computed based on the formulas mentioned in Table 2 and are

segregated into low (0–5 mm), medium (5–25 mm), and high (>25 mm) rainfall classes. The trends of POD and FAR for all 15 datasets computed against IMD are shown in Figures 5 and 6. APHRODITE under gauge-based category has high POD in low rainfall criteria, indicating accurate detection of rainfall events. Further, the values of POD followed a declining trend with the increase in rainfall intensity, that is, for medium and higher rainfall events. The FAR value was low (0.11), indicating that the number of false events detected by APHRODITE was less, and it further increased to 0.21 and 0.2 for medium and high rainfall events. Figure 5(a) shows that APHRODITE performed well in detecting rainfall, and it had moderate detection capability (POD = 0.5) up to 106 mm of rainfall intensity. SM2RAIN-CCI product under the satellite-only dataset category exhibited an intermediate ability in detecting rainfall in case of low rainfall events (POD = 0.6) and negligible (POD = 0.08) to no (POD = 0) detection performance in medium and high rainfall events,

TABLE 3: Continuous and categorical statistics of precipitation datasets compared to IMD.

	POD			FAR			CC	RMSE	Bias
	Low	Medium	High	Low	Medium	High			
APHRODITE	0.86	0.59	0.24	0.11	0.21	0.2	0.91	2.89	0.99
SM2RAIN-CCI	0.6	0.08	0	0.32	0.7	1	0.71	5.7	3
CHIRP	0.89	0.46	0.01	0.32	0.56	0.97	0.87	3.08	-0.1
CHIRPS 0.05	0.59	0.38	0.16	0.31	0.65	0.95	0.84	3.82	-0.78
CHIRPS 0.25	0.59	0.48	0.23	0.23	0.59	0.96	0.88	5.56	-2.1
GSMaP_Gauge_RNL	0.74	0.4	0.01	0.2	0.54	0.9	0.79	3.9	0.87
NCEP CFSR	0.93	0.55	0.03	0.3	0.53	0.95	0.86	3.33	-0.77
PGF 0.25	0.5	0.3	0.01	0.17	0.61	0.96	0.83	4.39	1.93
PGF 0.5	0.43	0.25	0.02	0.25	0.66	0.99	0.87	3.69	1.46
MSWEP	0.81	0.31	0.01	0.29	0.58	0.89	0.8	4.1	1.1
GPCP-CDR v1.3	0.45	0.12	0.01	0.33	0.76	1	0.73	5.26	2.44
GPCC v2018	0.79	0.68	0.29	0.22	0.63	0.97	0.88	10.13	-5.6
GPCC v7	0.77	0.52	0.07	0.21	0.59	0.96	0.88	3.62	-0.86
PERSIANN-CDR	0.37	0.1	0.03	0.44	0.88	1	0.34	6.42	2.78
TRMM 3B42 v7	0.58	0.38	0.14	0.58	0.38	0.14	0.9	2.82	0.48

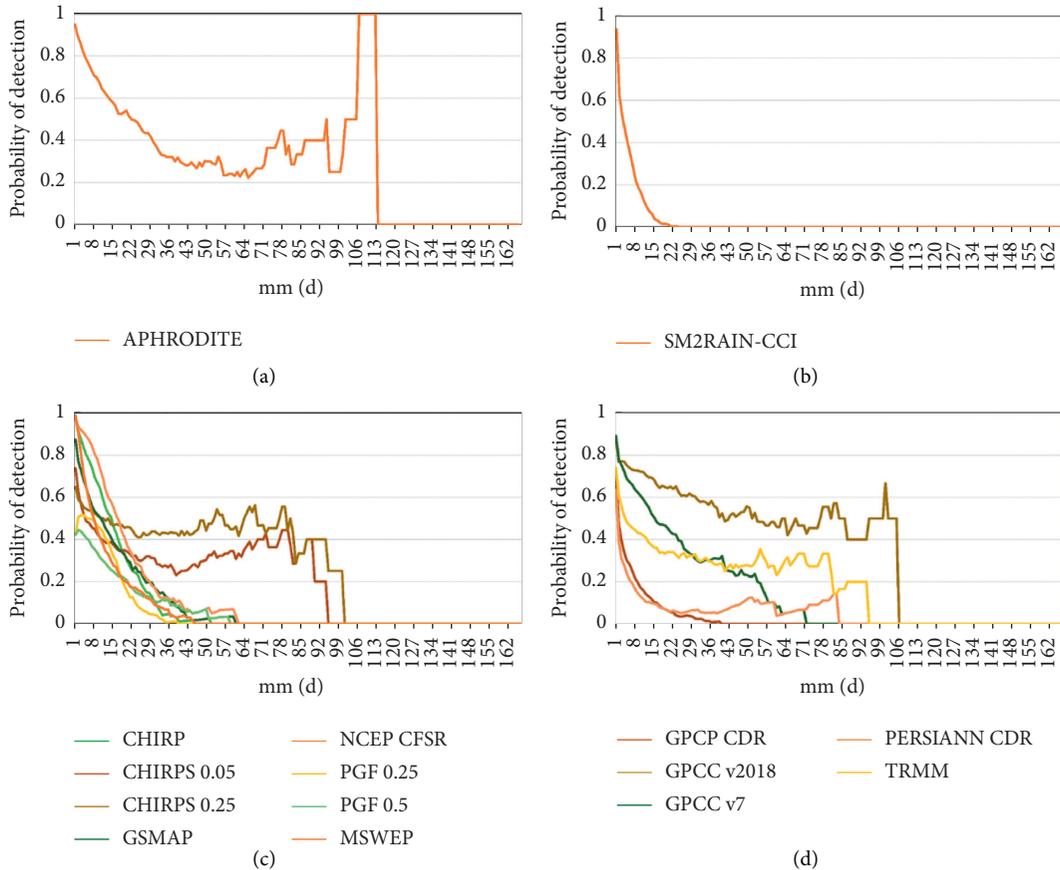


FIGURE 5: Line graphs indicating POD were plotted for (a) gauge-based product, (b) satellite-only product, (c) reanalysis products, and (d) gauge-adjusted products against IMD gridded dataset.

respectively. The satellite-based dataset (SM2RAIN-CCI) provided higher FAR values for low, medium, and high rainfall events ranging from 0.32 to 1, indicating that the dataset detected more “no-rain” or false events. Figure 5(b) shows that the SM2RAIN-CCI dataset has a clear declining

trend with the increase in precipitation intensity, and it detected until 22 mm of rainfall with low POD values. The FAR values plotted in Figure 6(b) exhibited an increasing trend, indicating that the dataset witnessed high false rainfall events.

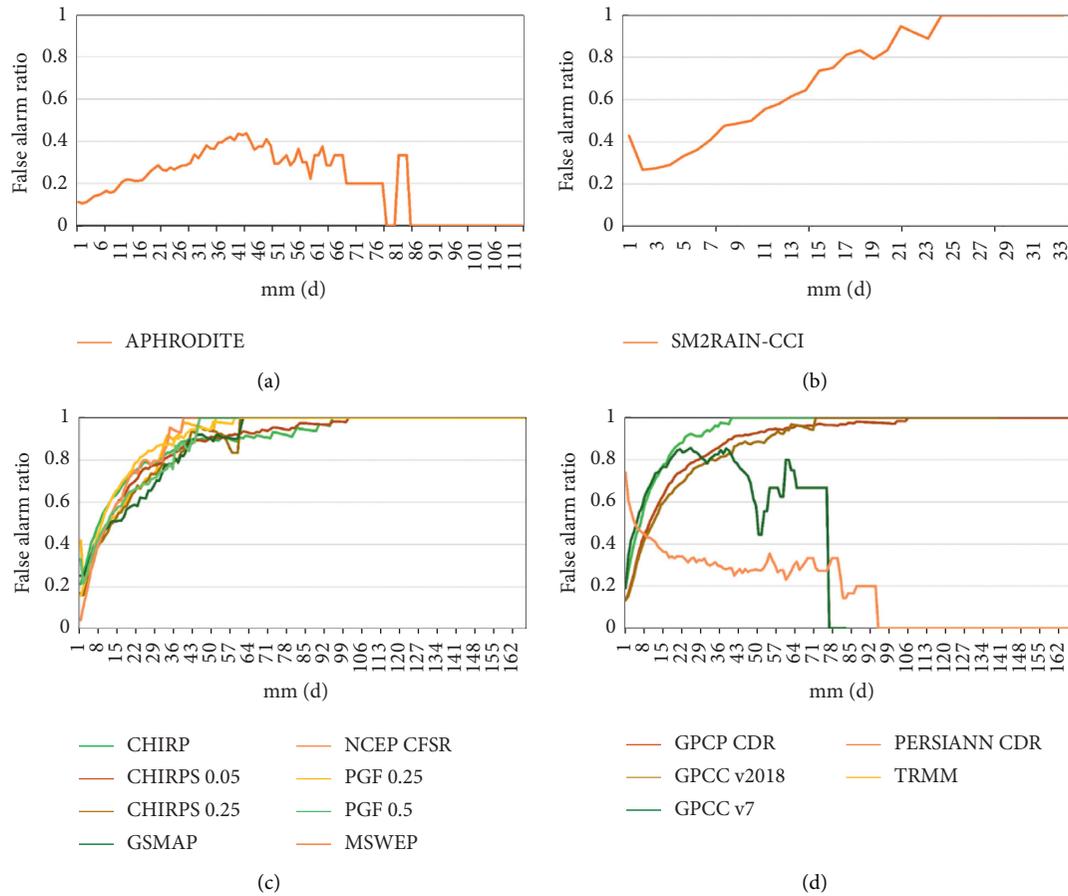


FIGURE 6: Line graphs indicating FAR were plotted for (a) gauge-based product, (b) satellite-only product, (c) reanalysis products, and (d) gauge-adjusted products against IMD gridded dataset.

(2) *From Reanalysis Datasets.* Under the reanalysis dataset category for low rainfall events (0–5 mm), NCEP-CFSR outperformed other rainfall datasets with POD and FAR values of 0.93 and 0.3, followed by the CHIRP dataset with similar performance in detecting low rainfall events (Figures 5(c) and 6(c)). Moderate performance was observed for medium rainfall events (6–25 mm) for both NCEP-CFSR and CHIRP with POD values of 0.55 and 0.46 and FAR values of 0.53 and 0.56, respectively. Poor performance in terms of POD and FAR was observed for NCEP-CFSR and CHIRP datasets in detecting high rainfall intensities (>25 mm). MSWEP and GSMaP\_Gauge\_RNL v6 datasets under the reanalysis category revealed higher POD values (0.81 and 0.74) and lower FAR values (0.29 and 0.18) for rain events between 0 and 5 mm, indicating better rainfall detection capabilities. Moderate performance was exhibited by CHIRPS 0.05° and CHIRPS 0.25° datasets in all three rainfall event classes (low, medium, and high). It can be observed from Figure 5(c) that only CHIRPS datasets had some ability to capture high rainfall events when compared to other datasets in the reanalysis category. All other datasets under the reanalysis category followed a similar trend with no rain detection in high or extreme rainfall events. PGF 0.25° dataset revealed moderate rainfall detection capability in low rainfall events with a POD of 0.5 and FAR of 0.17

(Figure 6(c)) but a poor performance in detecting medium and high rainfall events. PGF 0.5° dataset resulted in poor performance in all three rainfall events (low, medium, and high) under the reanalysis category.

(3) *From Gauge-Adjusted Datasets.* GPCCC v2018 and GPCCC v7 datasets manifested better performance in detecting low rainfall events. Under medium (5–25 mm) and high (>25 mm) rainfall events, GPCCC v2018 yielded moderate performance, whereas GPCCC v7 revealed poor performance. From Figure 5(d), it can be observed that GPCCC v2018 captured high rainfall events up to 106 mm/d with moderate detection capability. TRMM 3B42 v7 exhibited reasonable performance in detecting rainfall values across low, medium, and high rainfall events. Both GPCP-CDR v1.3 and PERSIANN-CDR exhibited poor performance with worse capabilities in all rainfall categories.

The POD values with 1 mm/day rainfall threshold exhibited that all datasets’ rainfall detection skill decreases as precipitation intensity increases, whereas the FAR metric shows an increase. This reveals that all the implemented datasets cannot capture the magnitude of extreme precipitation events accurately. The two categorical metrics (POD and FAR) showed that NCEP-CFSR has the best skill in detecting low rainfall events followed by CHIRP,

APHRODITE, MSWEP, and GPCC v2018 datasets. High rainfall events ( $>25$  mm) were best captured by GPCC v2018, CHIRPS 0.05° and CHIRPS 0.25°, and TRMM 3B42 v7 datasets.

**5.3. Performance Evaluation of Precipitation Datasets for Streamflow Simulations.** The values obtained from statistical coefficients, that is,  $R^2$ , N-S, and PBias, are represented in Table 4. The model was calibrated and validated at a monthly time step with a training period of 7 years ranging from 2002 to 2008 and a validation period of 4 years running from 2009 to 2012 with a warm-up period of 2 years.

**5.3.1. From the Results of N-S.** According to the ranges specified by [84] for streamflow simulations, IMD and APHRODITE in gauge-based category exhibited satisfactory performance in terms of estimating the acceptable magnitude of variances between observed and simulated datasets (N-S coefficient  $>0.65$ ). SM2RAIN-CCI dataset was left unsatisfactory in satellite-only listing. CHIRP v2.0 predicated excellent results supervised by GSMaP\_Gauge\_RNL v6 and CHIRPS 0.05° with good performance in the reanalysis section. TRMM 3B42 v7 outperformed in both gauge-adjusted and overall (out of 16 datasets) category in terms of producing less variance (N-S  $>0.75$ ) between the observed and simulated streamflow results.

**5.3.2. From the Results of PBias.** IMD dataset overestimated the flows with fewer biases and APHRODITE underestimated them with a more significant bias (PBias  $\geq 25\%$ ), indicating unsatisfactory performance in gauge-only category during calibration. SM2RAIN-CCI produced the highest mismatch between the simulated and observed streamflows and was left unsatisfactory. CHIRP v2.0, CHIRPS 0.05°, CHIRPS 0.25°, and GSMaP\_Gauge\_RNL v6 displayed excellent model simulations with fewer biases (PBias  $\leq 10\%$ ) in the reanalysis category. TRMM 3B42 v7 underestimated the streamflow with a lower bias of 7.1% and GPCC v2018 and GPCC v7 had an overestimation of  $-5.6\%$  and  $-9\%$  indicating very good performance (PBias  $\leq 10\%$ ) in gauge-adjusted category. PERSIANN-CDR and GPCP-CDR were underestimated with higher biases and were left unsatisfactory.

**5.3.3. From the Results of  $R^2$ .** IMD and APHRODITE obtained a moderate correlation ( $R^2 \geq 0.65$ ) (Figures 7(a) and 7(b)) under gauge-based category. Consistently SM2RAIN-CCI dataset resulted in unpropitious correlation (Figure 7(c)) under satellite-only listing. CHIRP v2.0 yielded a very good correlation between observed and simulated discharge values, which can be asserted from Figure 7(d). In contrast, scatterplots depicted in Figures 7(e) and 7(f) portray that GSMaP\_Gauge\_RNL v6 and CHIRPS 0.05° produced good performance ( $R^2 \geq 0.65$ ) with a moderate correlation. Remaining datasets in the reanalysis category (PGF 0.5° and 0.25°, MSWEP, CHIRPS 0.25°, and NCEP-CFSR) exhibited satisfactory analogue with  $R^2 \geq 0.5$  as

depicted in Figures 7(g)–7(k). Under gauge-adjusted classification, TRMM 3B42 v7 outperformed with excellent correlation ( $R^2 \geq 0.75$ , Figure 8(l)) ensued by GPCC v2018 and v7 with good analogue between observed and simulated streamflow values as shown in Figures 8(m) and 8(n).

It can be observed from Table 4 that IMD performed best in gauge-based category, whereas SM2RAIN-CCI performed the least overall. CHIRP v2.0 in the reanalysis category and TRMM 3B42 v7 in the gauge-adjusted category furnished excellent results. Overall, TRMM 3B42 v7 outperformed when compared with other 15 datasets in terms of  $R^2$  and N-S followed by CHIRP v2.0, GSMaP\_Gauge\_RNL v6, CHIRPS 0.05°, GPCC v7, GPCC v 2018, IMD gridded data, PGF 0.25°, APHRODITE, PGF 0.5°, MSWEP v1.2, CHIRPS 0.25°, PERSIANN-CDR, NCEP-CFSR, GPCP-CDR v1.3, and SM2RAIN-CCI datasets. To assert the overall performance of precipitation datasets, TRMM 3B42 v7 and CHIRP datasets during the calibration phase along with APHRODITE and PGF 0.25° during the validation phase furnished very good performance with high correlation, less variance, and fewer biases.

**5.4. Assessment of Extreme Flows.** Table 5 represents the peak and standard deviation values of observed and simulated streamflows. The peak value represents the maximum streamflow value in the observed/simulated dataset. In contrast, the standard deviation specifies the variance of the simulated dataset's mean with respect to the observed dataset. The model's simulated dataset should match all the peaks and low flows when represented in a graphical format to consider the model as effective. From Table 5, it can be seen that the observed streamflow from station data and simulated streamflow from the SWAT model have more significant differences in peak values during the calibration phase, contributing to the decrease in  $R^2$ , low correlation, and mismatch of peaks, which are represented in Figures 9(a)–9(e).

The SWAT model failed to capture the peaks correctly, since there is a larger deviation between the observed and simulated streamflow values (represented in Table 5) during the calibration period. From Table 5 and Figure 9(a) plotted for gauge-based datasets, it can be observed that IMD gridded data and APHRODITE have a more significant deviation and could not match the low flow events as well. SM2RAIN-CCI dataset portrayed in Figure 9(b) has more significant differences between observed and simulated peaks, resulting in unsatisfactory performance. From Figure 9(c) related to the reanalysis category, it can be seen that the CHIRP dataset matched the peak flows after 2004, whereas CHIRPS 0.05° was overestimated during the year 2005 and was underestimated for other years. In contrast, GSMaP\_Gauge\_RNL v6 tried to capture peak events compared to PGF 0.25° and that correctly captured the low flow events. MSWEP and PGF 0.5° datasets of the reanalysis section depicted in Figure 9(d) had more differences in peak and standard deviation values leading to lower  $R^2$  and higher PBias coefficients. In contrast, CHIRPS 0.25° was overestimated during the year 2005 and was underestimated for

TABLE 4: Statistical coefficient values for monthly runoff.

Precipitation dataset	$R^2$		N-S		PBias	
	Calibration (2002–2008)	Validation (2009–2012)	Calibration	Validation	Calibration	Validation
<i>Gauge-based</i>						
IMD	0.72	0.71	0.6	0.65	-4.2	20.2
APHRODITE	0.63	0.79	0.51	0.77	30.6	-3.5
<i>Satellite-only</i>						
SM2RAIN-CCI	0.48	0.4	0.18	0.16	58	57.8
<i>Reanalysis products</i>						
CHIRP v2.0	0.80	0.81	0.80	0.8	0.3	-9.7
GSMAP_Gauge_RNL v6	0.74	0.82	0.74	0.78	0.9	-17.5
CHIRPS 0.05° v2.0	0.73	0.83	0.73	0.75	0.3	-19.8
PGF 0.25 v2	0.66	0.82	0.63	0.8	11.7	11.8
PGF 0.5 v2	0.63	0.76	0.55	0.68	27	23.1
MSWEP v1.2	0.62	0.72	0.55	0.68	25.3	10.5
CHIRPS 0.25° v2.0	0.61	0.7	0.61	0.68	2.3	-7.1
NCEP-CFSR	0.61	0.61	0.52	0.6	-30	-1.3
<i>Gauge-adjusted</i>						
TRMM 3B42 v7	0.81	0.88	0.81	0.85	7.1	-12.1
GPCC v.2018	0.7	0.86	0.68	0.86	-5.6	1.8
GPCC v.7	0.73	0.83	0.7	0.77	-9	-27.4
PERSIANN-CDR	0.60	0.65	0.52	0.62	30.1	22.4
GPCP-CDR v1.3	0.38	0.67	0.35	0.63	22.7	14.2

other years and, lastly, the NCEP-CFSR dataset failed to capture both peaks and low flows correctly. The line diagram depicted for gauge-adjusted datasets in Figure 9(e) represents that TRMM 3B42 v7 performed very well and matched the peaks and low flows in all the years compared to other datasets. GPCC v2018 has less difference in peak and deviation values between observed and simulated periods when compared to PERSIANN-CDR and GPCP-CDR v1.3 under the gauge-adjusted category.

Coming to the validation phase, from Table 5 and Figure 10(a) plotted for gauge-based datasets, it can be seen that IMD and APHRODITE underestimated the results in the first two years and overestimated them during the last two years of the validation period with smaller biases and deviations. The SM2RAIN-CCI dataset flows plotted in Figure 10(b) failed to match the peak and low flow events except in 2010, resulting in poor performance and larger biases. CHIRP dataset under the reanalysis category (Figure 10(c)) could not match the peaks in all the years but has minor deviation when compared to observed data, which resulted in better values of N-S. GSMaP\_Gauge\_RNL v6 tried to reach the peak flows and resulted in overestimation, whereas PGF 0.25° underestimated the flows in all the years (Figure 10(c)). CHIRPS 0.05° classified under the reanalysis category overestimated the peaks in all the years and had more significant deviations resulting in a more substantial bias. MSWEP, NCEP-CFSR, and PGF 0.5° from Figure 10(d) underestimated the peaks in all years, whereas CHIRPS 0.25° underestimated the peaks and has fewer biases. TRMM 3B42 v7 and GPCC v7 from the gauge-adjusted category overestimated all the peak events. From Figure 10(e), it can be seen that though GPCC v2018, which tried to match the peaks, overestimated the peaks in 2010 and underestimated them in 2012, it produced better results in estimating peaks

when compared with other datasets. GPCP-CDR and PERSIANN-CDR categorized under gauge-adjusted datasets (Figure 10(e)) underestimated the peaks in all years with more substantial biases. From the extreme flow analysis, it was observed that CHIRP, TRMM 3B42 v7, GPCC v7, and APHRODITE datasets captured more peak flow events and hence can be further implemented for extreme event analysis.

**5.5. Discussion.** The discussion section was framed to compare the results and observations of the current study with other research around the world. To clearly segregate the comparisons, the discussion section is divided into 4 parts, that is, (i) from gauge-based dataset results, (ii) from satellite-only dataset results, (iii) from reanalysis dataset results, and (iv) from gauge-adjusted dataset results. Each section will tackle the results obtained from both statistical and hydrological analyses carried out in the current study.

**5.5.1. From Gauge-Based Dataset Results.** Gauge-based datasets (IMD and APHRODITE) developed by implementing different algorithms to interpolate precipitation values between sparsely spread gauging stations may contribute to lower accuracies. Many studies broached gauge-based dataset as standard and compared all the satellite datasets by calibrating the hydrological models with station or gauge-based gridded dataset and implementing those sensitive parameters for testing other satellite precipitation datasets [5, 85–87]. The performance of gauge-based datasets for hydrological modeling in our study was less effective, since IMD and APHRODITE did not produce better results for simulating streamflows than TRMM 3B42 v7 and CHIRP datasets. Similar results were

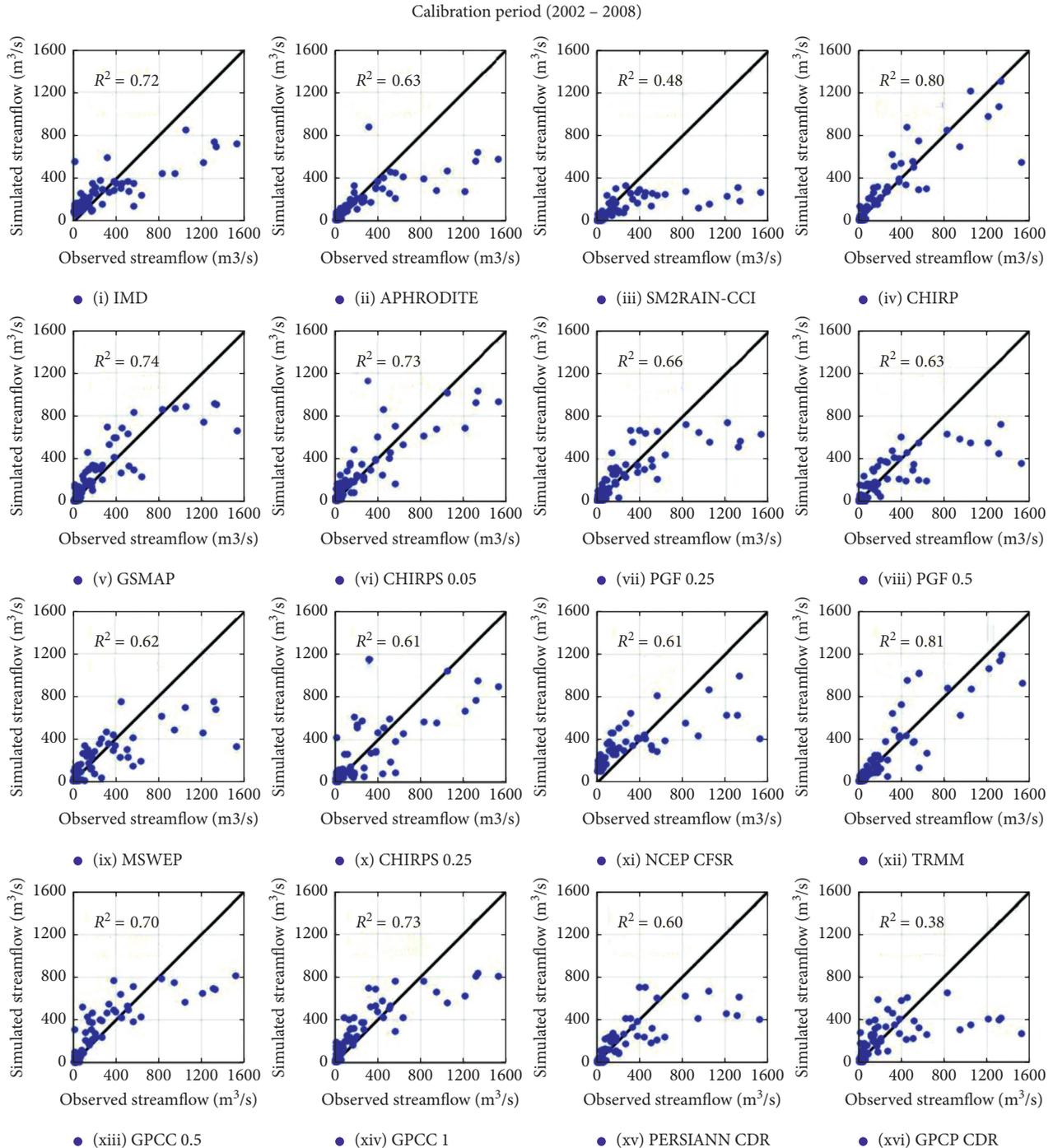


FIGURE 7: Scatterplots depicted for simulated versus observed runoff for (a) IMD, (b) APHRODITE, (c) SM2RAIN-CCI, (d) CHIRP, (e) GSMaP\_Gauge\_RNL v6, (f) CHIRPS 0.05°, (g) PGF 0.25°, (h) PGF 0.5°, (i) MSWEP, (j) CHIRPS 0.25°, (k) NCEP CFSR, (l) TRMM 3B42 v7, (m) GPCC v7, (n) GPCC v2018, (o) PERSIANN-CDR, and (p) GPCP-CDR v1.3 rainfall datasets during calibration phase.

published by [10, 42, 73], where satellite precipitation products produced better performance in simulating streamflows than using precipitation measured from gauging stations or from gauge-based gridded datasets. The lower accuracies of IMD and APHRODITE may be attributed to the high precision (using infrared and microwave data or highly accurate interpolation techniques for merging satellite data with gauge data) along with complete

spatial coverage of satellite datasets or usage of a different set of sensitive parameters for IMD, APHRODITE, and other datasets. The statistical results showed that CHIRP has less bias and better ability to detect low rainfall than APHRODITE (Table 5). This study illustrates to researchers that testing and calibrating each precipitation dataset are essential, since the standard dataset (which we assume as IMD) may not provide accurate results in every catchment.

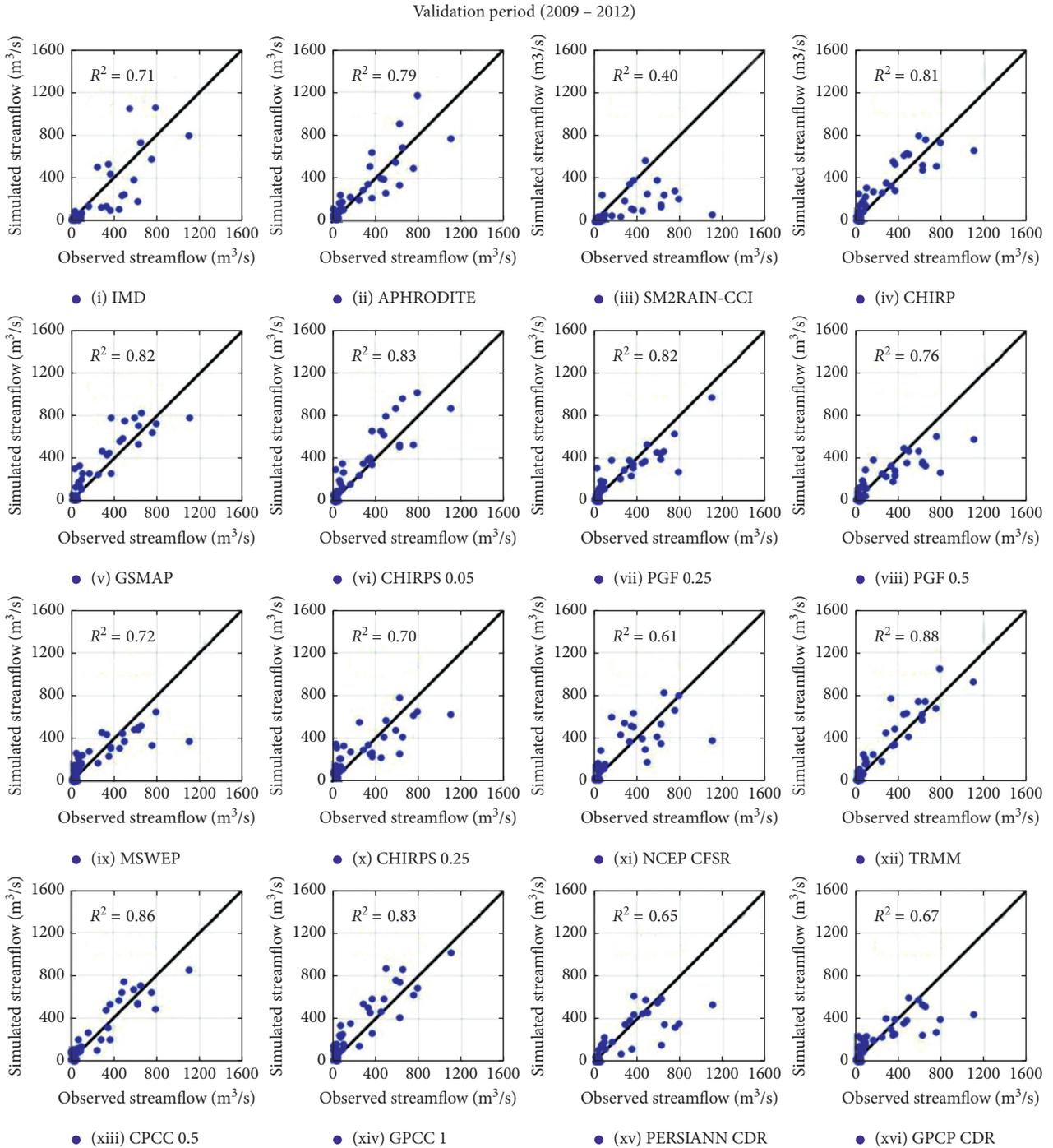


FIGURE 8: Scatterplots depicted for simulated versus observed runoff for (a) IMD, (b) APHRODITE, (c) SM2RAIN-CCI, (d) CHIRP, (e) GSMaP\_Gauge\_RNL v6, (f) CHIRPS 0.05°, (g) PGF 0.25°, (h) PGF 0.5°, (i) MSWEP, (j) CHIRPS 0.25°, (k) NCEP-CFSR, (l) TRMM 3B42 v7, (m) GPCP v7, (n) GPCP v2018, (o) PERSIANN-CDR, and (p) GPCP-CDR rainfall datasets during validation phase.

5.5.2. *From Satellite-Only Dataset Results.* We observed that SM2RAIN-CCI under satellite-only category furnished good correlation against IMD but failed to represent fewer RMSE and bias values from continuous statistical analyses, which are parallel to the results reported by [88]. They mentioned that the SM2RAIN-CCI dataset overestimates precipitation magnitudes at low- and high-altitude regions and

underestimates them in the medium-altitude regions. The SM2RAIN-CCI dataset (satellite-only category) employed in the present study exhibited more significant overestimation in detecting rainfall with substantial bias and a larger magnitude of errors. Though the dataset demonstrated overestimation in detecting rainfall, the streamflow simulations were underestimated. This discrepancy may be

TABLE 5: Streamflow characteristics during calibration and validation phases (Obs: observed; Sim: simulated).

Dataset	Calibration phase				Validation phase			
	Peak values (m <sup>3</sup> /s)		Standard deviation (m <sup>3</sup> /s)		Peak values (m <sup>3</sup> /s)		Standard deviation (m <sup>3</sup> /s)	
	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim
<i>Gauge-based</i>								
IMD	1523.13	853.5	333.88	166.50	1102.1	1077	264.25	271.53
APHRODITE		877.2		171.95		1158		262.60
<i>Satellite-only</i>								
SM2RAIN-CCI	1523.13	330.1	333.88	101.07	1102.1	565.6	264.25	130.79
<i>Reanalysis products</i>								
CHIRP	1523.13	1306	333.88	297.25	1102.1	792.1	264.25	250.53
CHIRPS 0.05°		1116		279.48		1012		297.69
CHIRPS 0.25°		1152		257.25		780.5		192.79
GSMaP		917.4		268.18		826.8		278.46
NCEP-CFSR		999.8		182.94		826.9		237.78
PGF 0.25°		741.3		216.25		966.2		208.24
PGF 0.5°		720.8		189.65		605.9		175.76
MSWEP		751.7		194.23		650.3		180.72
<i>Gauge-adjusted products</i>								
GPCP-CDR	1523.13	648.9	333.88	171.78	1102.1	592.9	264.25	172.30
GPCC v.7		836.1		234.89		1013		276.47
GPCC v.2018		817.1		239.53		851.5		253.34
PERSIANN-CDR		707.7		192.36		608.4		197.76
TRMM 3B42		1188		307.22		1038		291.62

attributed to the overcalibration of SWAT model sensitive parameters with the N-S coefficient as a primary objective function.

**5.5.3. From Reanalysis Dataset Results.** All the three CHPCLim datasets from the reanalysis category resulted in underestimation with moderate bias and RMSE for CHIRP and CHIRPS 0.05° datasets and high bias and RMSE for CHIRPS 0.25° datasets. The underestimation of CHPCLim datasets has been reported earlier by [42, 73, 82], where it was mentioned that the underestimation might be attributed to the infrared algorithm (which computes data from infrared region retrieved signals) used for CHPCLim dataset developments that implement fixed brightness temperature thresholds to distinguish between raining and nonraining clouds. The defined thresholds are usually too cold since the orographic precipitation occurring over the Western Ghats is warm, which may not produce much ice aloft, resulting in an underestimation of rainfall [89–93]. Coming to the categorical results, NCEP-CFSR outperformed other rainfall datasets followed by the CHIRP dataset with relatively similar performance in detecting low rainfall events. Similar results were published by [42, 79], where it was concluded that the NCEP-CFSR dataset outperformed other datasets in terms of categorical statistics in low rainfall events; however, it had a moderate-to-poor performance in terms of continuous statistical metrics.

MSWEP dataset (reanalysis category), which takes advantage of merging gauge, satellite, and reanalysis precipitation estimates, did not result in a better outcome in our study from both statistical and hydrological perspectives.

Though MSWEP has a better ability to detect low rainfall events, the magnitude of error and bias in the dataset was large, leading to the dataset's poor performance for simulating streamflows. The MSWEP datasets versions 1.2 and 2.0 were developed by considering several data sources that were tested at a global scale by [66], which concluded that MSWEP v2.0 outperforms other datasets, including MSWEP v1.2 in terms of statistical and hydrological criteria. Another study by [11] tested four satellite precipitation datasets in different climatic zones of peninsular Spain. It was concluded that MSWEP v2.0 performed best in a semiarid region and TRMM 3B42 v7 outperformed in other climatic regions (Oceanic climate, Galicia Variant, and Mediterranean climate). A study conducted by [94] concluded that TRMM 3B42 v7 and MSWEP v1.2 outperformed PERSIANN and CMORPH datasets in terms of statistical coefficients while simulating discharges and represented that TRMM 3B42 v7 and MSWEP v1.2 had similar tendency and correlation while simulating runoff. In the current study, MSWEP v1.2 was tested in a semihumid tropical region, and it was found that the performance of MSWEP was not very effective in the reanalysis category compared to CHIRP and CHIRPS products. Testing of MSWEP v1.2 product in different climatic and topographic zones is still required to understand the performance of MSWEP products as much literature was not found on this dataset. The performance of Princeton datasets (PGF 0.5° and PGF 0.25°) was moderate in both statistical and hydrological analyses because of high bias and less precipitation detection capabilities. From the results of PGF and CHIRPS SPPs that have multiple spatial resolutions, the datasets with finer resolution (PGF 0.25° and CHIRPS 0.05°) proved effective compared to coarse

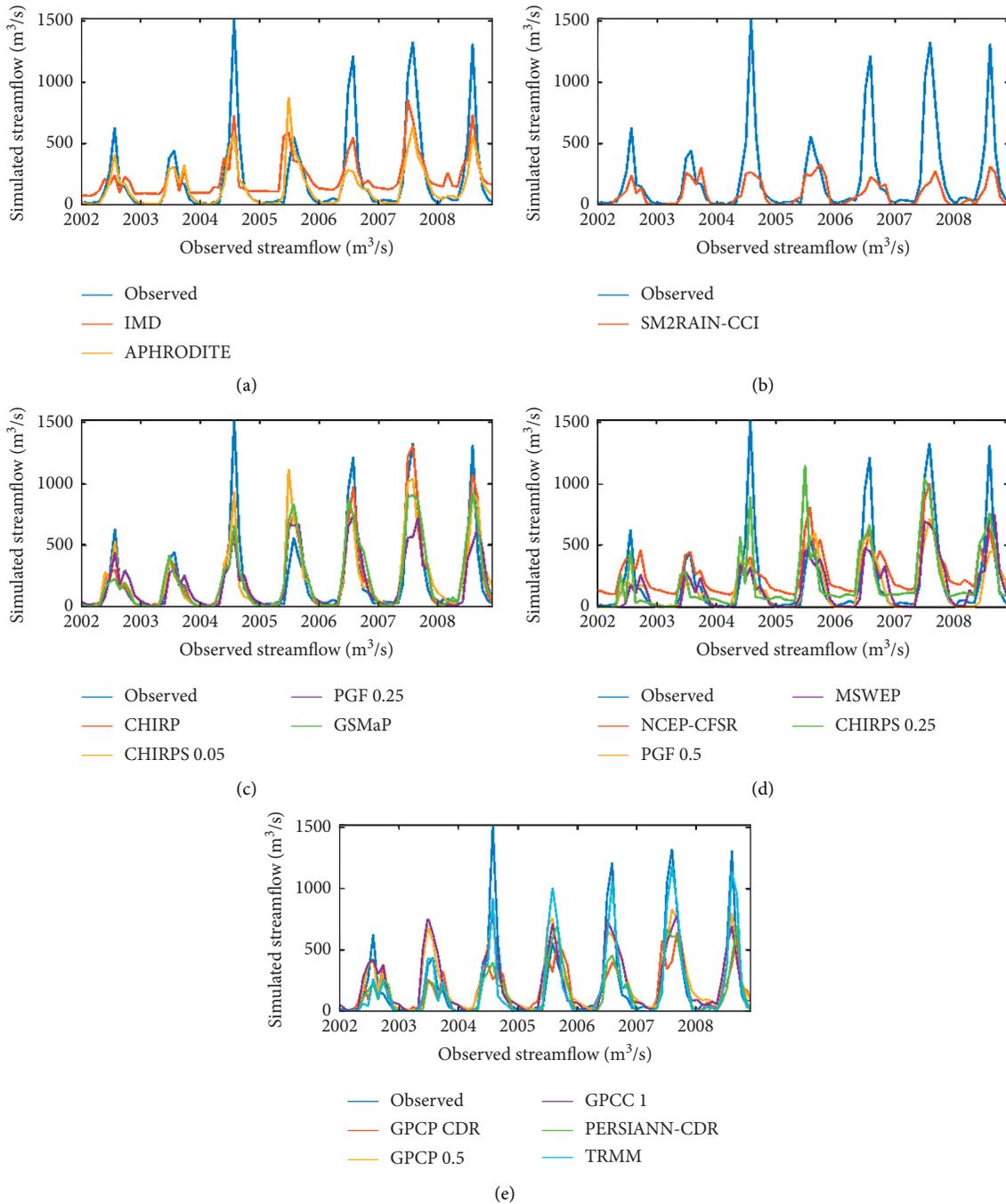


FIGURE 9: Monthly time series of simulated and observed runoff for observed and (a) gauge-based, (b) satellite-based, (c, d) reanalysis, and (e) gauge-adjusted datasets during the calibration phase.

resolution products (PGF 0.5° and CHIRPS 0.25°). One possible reason for the better performance of fine scale products can be attributed to the size of the catchment considered in the study. As the study area has a catchment area of 7778 km<sup>2</sup>, coarse resolution products (0.5° and 1° spatial resolution) might have 20–25 rainfall grid cells over the basin, whereas finer resolution (0.05° and 0.25° spatial resolution) products might accommodate 70–90 grids. Lesser number of grid cells lacks the ability to capture the heterogeneity of rainfall in a mountainous basin, thus

leading to poorer performance when compared to fine scale SPPs. However, the interpolation algorithm, quality, resolution, time period, and blending procedures of input parameters will also play a pivotal role in proving the effectiveness of the dataset.

5.5.4. From Gauge-Adjusted Dataset Results. Poor performance and overestimation were observed for PERSIANN-CDR under gauge-adjusted category and this may be

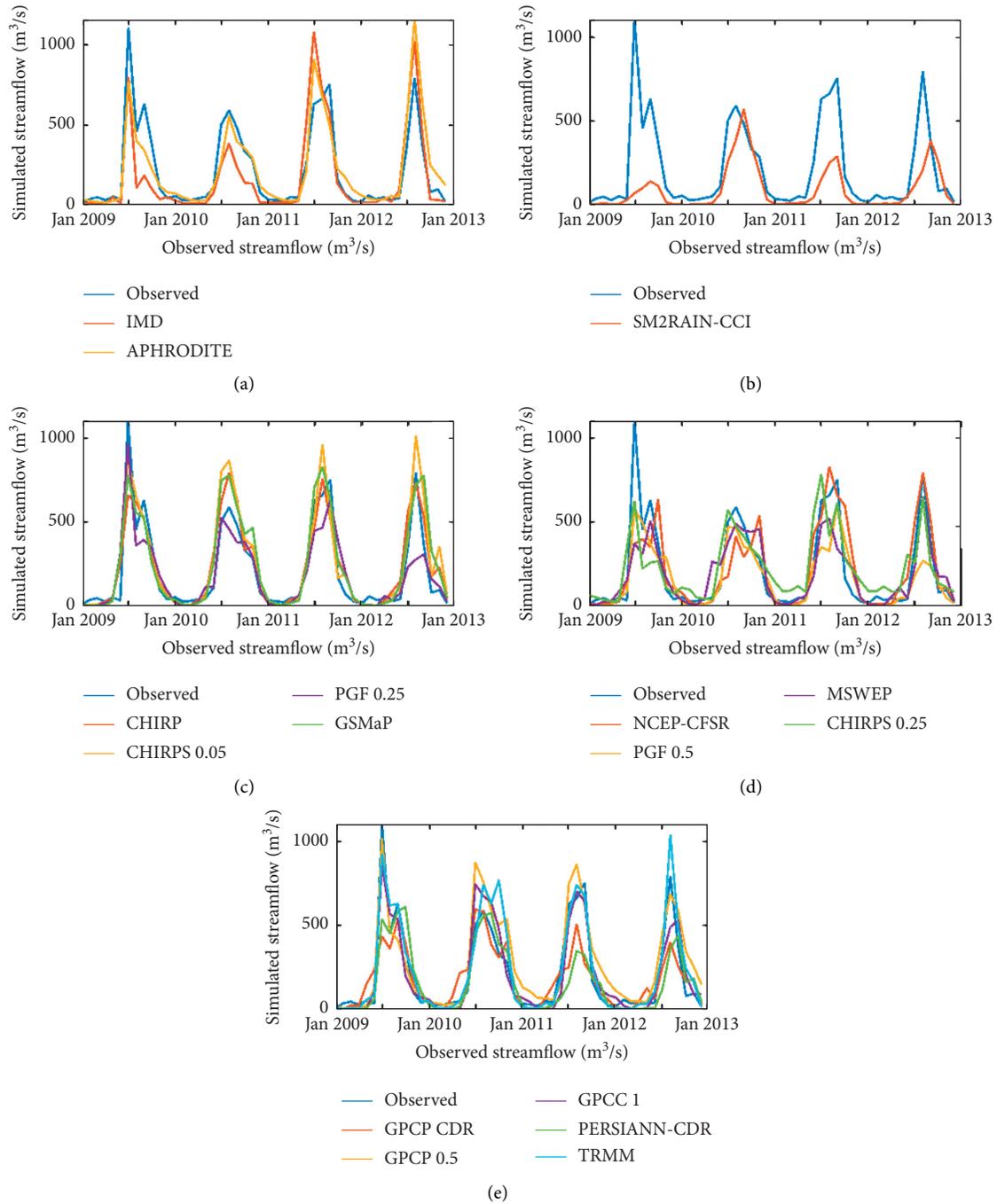


FIGURE 10: Monthly time series of simulated and observed runoff for observed and (a) gauge-based, (b) satellite-only, ((c) & (d)) reanalysis, and (e) gauge-adjusted datasets during the validation phase.

attributed to the inadequate training of ANN over parts of the world other than the United States because PERSIANN-CDR dataset is competently trained only over the United States [83, 95]. Though CHIRPS afforded a good performance, TRMM 3B42 v7 outperformed those results, which may be due to the implementation of combined infrared and microwave data in the development of TRMM 3B42 v7, whereas only infrared signals are used for the development of CHIRPS products. The biases were less for the TRMM 3B42 v7 product, which may be due to the inclusion of the

inverse-error-variance-weighting algorithm and calibration with a gauge analysis product of GPCC during development. From the hydrologic evaluations in the present study, it can be apprehended that TRMM 3B42 v7 also provided better results than the CHIRP dataset in simulating streamflows, whereas CHIRP performed better than TRMM 3B42 v7 in detecting rainfall, which is evident through statistical analysis results. Similar results were cited by [82, 86, 96], where the applicability of these products in a variety of catchments having different topographic and climatic

patterns was tested using hydrological modeling. It was concluded that though TRMM 3B42 v7 and CHIRP have certain biases, they outperformed other satellite precipitation datasets.

From the overall analysis, it can be observed that though the satellite products exhibited higher correlations at monthly time step in the statistical analysis category, satellite product performance varied significantly in simulating streamflows. The CHIRP dataset proved effective in both hydrological and statistical analyses. In contrast, the NCEP-CFSR dataset, which detected rainfall effectively, did not perform better in simulating streamflows and TRMM 3B42 v7, which has less precipitation detection capability, outperformed in streamflow simulations. These discrepancies can be attributed to the bias and error magnitudes of a particular product. TRMM 3B42 v7 and CHIRP, which have less bias, performed well, whereas NCEP-CFSR, PGF, GPCP-CDR v1.3, and PERSIANN CDR, which have more bias, exhibited moderate performance in streamflow simulations. Our results are in parallel with the conclusions made by [79], stating that the bias of rainfall products determines the accuracy of runoff simulations by a model. The reason specified for the conclusion (streamflow simulations are affected by the bias of a product) as per [16, 42] is related to the nonlinearity of the hydrologic process, where moderate rainfall bias can be transmuted into large PBias in discharge simulations. In correspondence to the above conclusion, the RMSE and bias values of SM2RAIN-CCI and GPCP-CDR v1.3 might have adversely affected the hydrologic performance and contributed to unsatisfactory discharge simulations.

## 6. Conclusions

This study compared sixteen rainfall datasets belonging to four categories (gauge-based, satellite-based, reanalysis, and gauge-adjusted products) in a mountainous tropical catchment of Karnataka, that is, Tungabhadra. Tungabhadra, with hilly terrain and varying topography, was selected to analyze the (i) capability of satellite precipitation datasets in detecting rainfall through statistical analysis and (ii) hydrological performance of these precipitation datasets in simulating streamflow when compared against observed streamflow data. Since the watershed is situated at higher elevations with mountainous terrain, the density of rain gauges in and around the watershed is very low, mandating us to test satellite precipitation products' performance. Hence, the main aim of the present study is to assess the (i) capability of precipitation datasets in detecting rainfall and (ii) suitability of sixteen rainfall datasets for hydrological modeling tested in terms of statistical coefficients ( $R^2$ , N-S, and PBias). From the results of continuous statistical metrics, APHRODITE furnished very good results in detecting rainfall with less bias and error magnitudes when compared against the IMD dataset. CHIRP, NCEP-CFSR, TRMM 3B42 v7, GSMaP\_Gauge\_RNL v6, and MSWEP datasets exhibited good-to-moderate performances at a monthly time step. From the results of categorical statistical metrics, it was revealed that NCEP-CFSR has the best skill in detecting low

rainfall events followed by CHIRP, APHRODITE, MSWEP, and GPCC v2018 datasets. High rainfall events (>25 mm) were tried to be captured by GPCC v2018, CHIRPS 0.05° and 0.25°, and TRMM 3B42 v7 datasets. The objective of testing the suitability of these precipitation datasets was successfully achieved as three datasets under the reanalysis category (CHIRP, CHIRPS 0.05°, and GSMaP\_Gauge\_RNL v6) and three datasets under gauge-adjusted category (TRMM 3B42 v7, GPCC v7, and GPCC v2018) are under ranges specified by [84] in simulating streamflow using a hydrological model, that is, SWAT, for Tungabhadra river basin, India. The newly released GPCC v.2018 proved effective when compared to GPCP-CDR v1.3. The LULC classification using a maximum likelihood algorithm produced an accuracy of 85.94%, indicating that the classification has achieved good accuracy. Sensitive parameters were obtained from previous literature and are used in different sets based on the sensitivity to each precipitation dataset in SWAT CUP for valid calibration and validation results. The present study showed that changes in input parameters affect the output results. The quality, resolution, time period, and blending procedures of input parameters will also play a pivotal role in proving the effectiveness of the dataset and achieving accurate results. It was shown from CHIRPS and PGF datasets that finer resolutions provide better results compared to coarser resolutions in case of the same type of dataset, which was developed by implementing the same algorithm and same input parameters. From both statistical and hydrological performance outputs, it can be concluded that rainfall product bias determines the accuracy of hydrological model runoff simulations. This is mainly because the bias of rainfall products will be transmuted into larger PBias during runoff simulations using a hydrological model. Out of 16 datasets, one can use TRMM 3B42 v7, CHIRP, CHIRPS 0.05°, and GSMaP\_Gauge\_RNL v6 datasets for hydrological modeling, climate change studies, and other research in similar topographic and climatic watersheds as they achieved overall very good performance.

The highlights of this paper are listed as follows:

- (i) Eight out of sixteen datasets yielded good results even without bias correction
- (ii) TRMM 3B42 v7 proved best for streamflow modeling and APHRODITE was the best in detecting rainfall
- (iii) The reliability and applicability of recently released GPCP-CDR v1.3 and GPCC v2018 datasets were explored

## Data Availability

All the datasets used in this study are freely available online. The IMD gridded data can be accessed from <http://www.imdpune.gov.in>, APHRODITE from <http://aphrodite.st.hirosaki-u.ac.jp>, SM2RAIN-CCI from <https://zenodo.org/record/1305021>, CHIRP from <ftp://ftp.chg.ucsb.edu/pub/org/chg/products/>, GSMaP from <https://sharaku.eorc.jaxa.jp/GSMaP/>, NCEP-CFSR from <https://globalweather.tamu.edu/>,

PGF from <http://hydrology.princeton.edu/data/pgf>, MSWEP from <http://www.gloh2o.org/>, GPCP-CDR from [http://eagle1.umd.edu/GPCP\\_CDR](http://eagle1.umd.edu/GPCP_CDR), PERSIANN CDR from <http://chrsdata.eng.uci.edu/>, and TRMM from <https://disc.gsfc.nasa.gov/>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors thank Mr. Kolluru Srinivas (Research Scholar, IIT Bombay) for his support in data extraction and Dr. Archana Mohite (IIT Kharagpur) for her assistance in performing statistical analysis. The authors gratefully thank all institutes, organizations, and developers of precipitation datasets for providing their products freely.

## Supplementary Materials

Supplementary Table 1: sensitivity parameters for SWAT CUP. Supplementary Figure 1: soil and land use map of the study area. Supplementary Information: overview on satellite precipitation datasets. (*Supplementary Materials*)

## References

- [1] C. Kidd, A. Becker, G. J. Huffman et al., "So, how much of the earth's surface is covered by rain gauges?" *Bulletin of the American Meteorological Society*, vol. 98, no. 1, pp. 69–78, 2017.
- [2] E. Sharifi, R. Steinacker, and B. Saghaian, "Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: preliminary results," *Remote Sensing*, vol. 8, no. 2, p. 135, 2016.
- [3] M. L. Tan, P. W. Gassman, X. Yang, and J. Haywood, "A review of SWAT applications, performance and future needs for simulation of hydro-climatic extremes," *Advances in Water Resources*, vol. 143, p. 103662, 2020.
- [4] A. K. Shukla, C. S. P. Ojha, and R. D. Garg, "Geo-spatial approach for estimation of precipitation over the upper Ganga river basin (UGRB)," vol. 9, 2016.
- [5] N. Ghodichore, R. Vinnarasi, C. T. Dhanya, and S. B. Roy, "Reliability of reanalyses products in simulating precipitation and temperature characteristics over India," *Journal of Earth System Science*, vol. 127, no. 8, 2018.
- [6] S. Prakash, A. K. Mitra, and D. S. Pai, "Comparing two high-resolution gauge-adjusted multisatellite rainfall products over India for the southwest monsoon period," *Meteorological Applications*, vol. 22, no. 3, pp. 679–688, 2015.
- [7] V. Kolluru, S. Kolluru, and P. Konkathi, "Evaluation and integration of reanalysis rainfall products under contrasting climatic conditions in India," *Atmospheric Research*, vol. 246, Article ID 105121, 2020.
- [8] E. Sharifi, J. Eitzinger, and W. Dorigo, "Performance of the state-of-the-art gridded precipitation products over mountainous terrain: a regional study over Austria," *Remote Sensing*, vol. 11, no. 17, pp. 2018–2020, 2019.
- [9] K. Pakoksung and M. Takagi, "Effect of satellite based rainfall products on river basin responses of runoff simulation on flood event," *Modeling Earth Systems and Environment*, vol. 2, no. 3, p. 143, 2016.
- [10] F. Satgé, D. Ruelland, M.-P. Bonnet, J. Molina, and R. Pillco, "Consistency of satellite precipitation estimates in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region," *Hydrology and Earth System Sciences*, pp. 1–41, 2018.
- [11] J. Senent-Aparicio, A. López-Ballesteros, J. Pérez-Sánchez, F. Segura-Méndez, and D. Pulido-Velazquez, "Using multiple monthly water balance models to evaluate gridded precipitation products over peninsular Spain," *Remote Sensing*, vol. 10, no. 6, p. 922, 2018.
- [12] S. Aminyavari, B. Saghaian, and E. Sharifi, "Assessment of precipitation estimation from the NWP models and satellite products for the spring 2019 severe floods in Iran," *Remote Sensing*, vol. 11, no. 23, p. 2741, 2019.
- [13] G. Artan, H. Gadain, J. L. Smith, K. Asante, C. J. Bandaragoda, and J. P. Verdin, "Adequacy of satellite derived rainfall data for stream flow modeling," *Natural Hazards*, vol. 43, no. 2, pp. 167–185, 2007.
- [14] B. Collischonn, W. Collischonn, and C. E. M. Tucci, "Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates," *Journal of Hydrology*, vol. 360, no. 1–4, pp. 207–216, 2008.
- [15] F. Su, Y. Hong, and D. P. Lettenmaier, "Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin," *Journal of Hydrometeorology*, vol. 9, no. 4, pp. 622–640, 2008.
- [16] K. Venkatesh, K. Preethi, and H. Ramesh, "Evaluating the effects of forest fire on water balance using fire susceptibility maps," *Ecological Indicators*, vol. 110, Article ID 105856, 2020.
- [17] A. Behrangi, B. Khakbaz, T. C. Jaw, A. AghaKouchak, K. Hsu, and S. Sorooshian, "Hydrologic evaluation of satellite precipitation products over a mid-size basin," *Journal of Hydrology*, vol. 397, no. 3–4, pp. 225–237, 2011.
- [18] S. Camici, L. Ciabatta, C. Massari, and L. Brocca, "How reliable are satellite precipitation estimates for driving hydrological models: a verification study over the Mediterranean area," *Journal of Hydrology*, vol. 563, pp. 950–961, 2018.
- [19] A. S. Falck, V. Maggioni, J. Tomasella, D. A. Vila, and F. L. R. Diniz, "Propagation of satellite precipitation uncertainties through a distributed hydrologic model: a case study in the Tocantins-Araguaia basin in Brazil," *Journal of Hydrology*, vol. 527, pp. 943–957, 2015.
- [20] V. Kolluru, S. Kolluru, N. Wagle, and T. D. Acharya, "Secondary precipitation estimate merging using machine learning: development and evaluation over Krishna river basin, India," *Remote Sensing*, vol. 12, no. 18, p. 3013, 2020.
- [21] M. Zambrano-Bigiarini, A. Nauditt, C. Birkel, K. Verbist, and L. Ribbe, "Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile," *Hydrology and Earth System Sciences*, vol. 21, no. 2, pp. 1295–1320, 2017.
- [22] S. Moazami, S. Golian, M. R. Kavianpour, and Y. Hong, "Comparison of PERSIANN and V7 TRMM multi-satellite precipitation analysis (TMPA) products with rain gauge data over Iran," *International Journal of Remote Sensing*, vol. 34, no. 22, pp. 8156–8171, 2013.
- [23] N. Voisin, A. W. Wood, and D. P. Lettenmaier, "Evaluation of precipitation products for global hydrological prediction," *Journal of Hydrometeorology*, vol. 9, no. 3, pp. 388–407, 2008.
- [24] L. Li, C. S. Ngongondo, C.-Y. Xu, and L. Gong, "Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa," *Hydrology Research*, vol. 44, no. 5, pp. 770–788, 2013.

- [25] H. E. Beck, M. Pan, T. Roy et al., "Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS," *Hydrology and Earth System Sciences*, pp. 1–23, 2018.
- [26] G. R. C. Essou, R. Arsenault, and F. P. Brissette, "Comparison of climate datasets for lumped hydrological modeling over the continental United States," *Journal of Hydrology*, vol. 537, pp. 334–345, 2016.
- [27] S. Wang, S. Liu, X. Mo et al., "Evaluation of remotely sensed precipitation and its performance for streamflow simulations in basins of the Southeast Tibetan Plateau," *Journal of Hydrometeorology*, vol. 16, no. 6, pp. 2577–2594, 2015.
- [28] C. Casse, M. Gosset, C. Peugeot et al., "Potential of satellite rainfall products to predict Niger River flood events in Niamey," *Atmospheric Research*, vol. 163, pp. 162–176, 2015.
- [29] G. Tang, Z. Zeng, D. Long et al., "Statistical and hydrological comparisons between TRMM and GPM Level-3 products over a midlatitude Basin: is day-1 IMERG a good successor for TMPA 3B42V7?" *Journal of Hydrometeorology*, vol. 17, no. 1, pp. 121–137, 2016.
- [30] S. Stisen and I. Sandholt, "Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling," *Hydrological Processes*, vol. 24, no. 7, pp. 879–891, 2010.
- [31] S. Hussain, X. Song, G. Ren, I. Hussain, D. Han, and M. H. Zaman, "Evaluation of gridded precipitation data in the Hindu Kush-Karakoram-Himalaya mountainous area," *Hydrological Sciences Journal*, vol. 62, no. 14, pp. 2393–2405, 2017.
- [32] Y. Yang, G. Wang, L. Wang, J. Yu, and Z. Xu, "Evaluation of gridded precipitation data for driving SWAT model in area upstream of three gorges reservoir," *PLoS One*, vol. 9, no. 11, Article ID e112725, 2014.
- [33] A. Shukla, C. Ojha, R. Singh, L. Pal, and D. Fu, "Evaluation of TRMM precipitation dataset over Himalayan catchment: the upper Ganga basin, India," *Water*, vol. 11, no. 3, p. 613, 2019.
- [34] H. E. Beck, N. Vergopolan, M. Pan et al., "Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling," *Hydrology and Earth System Sciences*, vol. 21, no. 12, pp. 6201–6217, 2017.
- [35] M. Mazzoleni, L. Brandimarte, and A. Amaranto, "Evaluating precipitation datasets for large-scale distributed hydrological modelling," *Journal of Hydrology*, vol. 578, p. 124076, 2019.
- [36] M. B. Abbott, J. C. Bathurst, J. A. Cunge, P. E. O. Connell, and J. Rasmussen, "An introduction to the European hydrological System - systeme Hydrologique Europeen, "SHE", 2: structure of a physically-based, distributed modelling system," *Journal of Hydrology*, vol. 87, no. 1-2, pp. 61–77, 1986.
- [37] C. Rasmussen, D. G. Tarboton, and R. Woods, "Application of TOPNET in the distributed model intercomparison project," *Journal of Hydrology*, vol. 298, no. 1-4, pp. 178–201, 2004.
- [38] Y. Qu and C. J. Duffy, "A semidiscrete finite volume formulation for multiprocess watershed simulation," *Water Resources Research*, vol. 43, no. 8, pp. 1–18, 2007.
- [39] J. Wang, Y. Hong, L. Li et al., "The coupled routing and excess storage (CREST) distributed hydrological model," *Hydrological Sciences Journal*, vol. 56, no. 1, pp. 84–98, 2011.
- [40] D. Lohmann, E. Raschke, B. Nijssen, and D. P. Lettenmaier, "Regional scale hydrology: I. formulation of the VIC-2L model coupled to a routing model," *Hydrological Sciences Journal*, vol. 43, no. 1, pp. 131–141, 1998.
- [41] R. Kommu, S. Kundapura, and V. Kolluru, "A statistical approach for comparison of secondary precipitation products," in *Trends in Civil Engineering and Challenges for Sustainability; Lecture Notes in Civil Engineering*, pp. 753–763, 2021, ISBN 9789811568275.
- [42] M. Musie, S. Sen, and P. Srivastava, "Comparison and evaluation of gridded precipitation datasets for streamflow simulation in data scarce watersheds of Ethiopia," *Journal of Hydrology*, vol. 579, Article ID 124168, 2019.
- [43] M. K. Shrestha, F. Recknagel, J. Frizenschaf, and W. Meyer, "Assessing SWAT models based on single and multi-site calibration for the simulation of flow and nutrient loads in the semi-arid Onkaparinga catchment in South Australia," *Agricultural Water Management*, vol. 175, pp. 61–71, 2016.
- [44] S. Kundu, D. Khare, and A. Mondal, "Individual and combined impacts of future climate and land use changes on the water balance," *Ecological Engineering*, vol. 105, pp. 42–57, 2017.
- [45] K. Venkatesh, H. Ramesh, and P. Das, "Modelling stream flow and soil erosion response considering varied land practices in a cascading river basin," *Journal of Environmental Management*, vol. 264, p. 110448, 2020.
- [46] M. S. Kumar, V. Kolluru, S. B. Gowthami et al., "Monitoring land use and land cover changes in coastal Karnataka," in *Trends in Civil Engineering and Challenges for Sustainability; Lecture Notes in Civil Engineering*, pp. 785–795, 2021, ISBN 9789811568275.
- [47] K. Venkatesh and H. Ramesh, "Impact of land use land cover change on run off generation in Tungabhadra river basin," in *Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Delft, Netherlands, October 2018.
- [48] T. M. Sharannya, A. Mudbhalkar, and A. Mahesha, "Assessing climate change impacts on river hydrology-a case study in the Western Ghats of India," *Journal of Earth System Science*, vol. 127, no. 6, pp. 1–11, 2018.
- [49] P. P. Saha, K. Zeleke, and M. Hafeez, "Streamflow modeling in a fluctuant climate using SWAT: Yass River catchment in south eastern Australia," *Environmental Earth Sciences*, vol. 71, no. 12, pp. 5241–5254, 2014.
- [50] P. Phomcha, P. Wirojanagud, T. Vangpaisal, and T. Thaveevouthti, "Modeling the impacts of alternative soil conservation practices for an agricultural watershed with the SWAT model," *Procedia Engineering*, vol. 32, pp. 1205–1213, 2012.
- [51] Y. Kheereemangkla, R. P. Shrestha, S. Shrestha, and D. Jourdain, "Modeling hydrologic responses to land management scenarios for the Chi River Sub-basin Part II, Northeast Thailand," *Environmental Earth Sciences*, vol. 75, no. 9, pp. 1–16, 2016.
- [52] S. Shrestha, B. Bhatta, M. Shrestha, and P. K. Shrestha, "Integrated assessment of the climate and landuse change impact on hydrology and water quality in the Songkhram River Basin, Thailand," *Science of the Total Environment*, vol. 643, pp. 1610–1622, 2018.
- [53] A. Malagó, F. Bouraoui, O. Vigiak, B. Grizzetti, and M. Pastori, "Modelling water and nutrient fluxes in the Danube river basin with SWAT," *Science of the Total Environment*, vol. 603–604, pp. 196–218, 2017.
- [54] K. Chhuon, E. Herrera, and K. Nadaoka, "Application of integrated hydrologic and river basin management modeling for the optimal development of a multi-purpose reservoir project," *Water Resources Management*, vol. 30, no. 9, pp. 3143–3157, 2016.
- [55] M. E. Arias, T. Piman, H. Lauri, T. A. Cochrane, and M. Kummu, "Dams on Mekong tributaries as significant contributors of hydrological alterations to the Tonle Sap

- Floodplain in Cambodia,” *Hydrology and Earth System Sciences*, vol. 18, no. 12, pp. 5303–5315, 2014.
- [56] M. L. Tan, P. W. Gassman, R. Srinivasan, J. G. Arnold, and X. Yang, “A review of SWAT studies in Southeast Asia: applications, challenges and future directions,” *Water*, vol. 11, no. 5, pp. 914–925, 2019.
- [57] J. Yang, I. Maximov, R. Srinivasan et al., “Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT,” *Journal of Hydrology*, vol. 333, pp. 413–430, 2006.
- [58] R. Meenu, S. Rehana, and P. P. Mujumdar, “Assessment of hydrologic impacts of climate change in Tunga-Bhadra river basin, India with HEC-HMS and SDSM,” *Hydrological Processes*, vol. 27, no. 11, pp. 1572–1589, 2013.
- [59] D. S. Pai, L. Sridhar, M. Rajeevan, O. P. Sreejith, N. S. Satbhai, and B. Mukhopadhyay, “Development of a new high spatial resolution ( $0.25^\circ \times 0.25^\circ$ ) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region,” *Mausam*, vol. 65, pp. 1–18, 2014.
- [60] A. Yatagai, K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, “Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges,” *Bulletin of the American Meteorological Society*, vol. 93, no. 9, pp. 1401–1415, 2012.
- [61] L. Ciabatta, C. Massari, L. Brocca et al., “SM2RAIN-CCI: a new global long-term rainfall data set derived from ESA CCI soil moisture,” *Earth System Science Data*, vol. 10, no. 1, pp. 267–280, 2018.
- [62] C. Funk, P. Peterson, M. Landsfeld et al., “The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes,” *Scientific Data*, vol. 2, no. 1, pp. 1–21, 2015.
- [63] K. Okamoto, T. Ushio, T. Iguchi, N. Takahashi, and K. Iwanami, “The global satellite mapping of precipitation (GSMaP) project,” in *Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium*, vol. 5, pp. 3414–3416, Seoul, Korea, July 2005.
- [64] S. Saha and H.-L. P. Shrinivas Moorthi, “The NCEP climate forecast system reanalysis,” *American Meteorological Society*, vol. 16, pp. 1015–1058, 2008.
- [65] J. Sheffield, G. Goteti, and E. F. Wood, “Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling,” *Journal of Climate*, vol. 19, no. 13, pp. 3088–3111, 2006.
- [66] H. E. Beck, A. I. J. M. Van Dijk, V. Levizzani et al., “MSWEP: 3-hourly  $0.25^\circ$  global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data,” *Hydrology and Earth System Sciences*, vol. 21, no. 1, pp. 589–615, 2017.
- [67] G. J. Huffman, R. F. Adler, M. M. Morrissey et al., “Global precipitation at one-degree daily resolution from multi-satellite observations,” *Journal of Hydrometeorology*, vol. 2, no. 1, pp. 36–50, 2001.
- [68] U. Schneider, A. Becker, M. Ziese, and B. Rudolf, “Global precipitation analysis products of the GPCP, Internet Publikation,” 2015.
- [69] D. Zhang, X. Liu, P. Bai, and X.-H. Li, “Suitability of satellite-based precipitation products for water balance simulations using multiple observations in a humid catchment,” *Remote Sensing*, vol. 11, no. 2, p. 151, 2019.
- [70] Z. Wu, Z. Xu, F. Wang et al., “Hydrologic evaluation of multi-source satellite precipitation products for the upper Huaihe river basin, China,” *Remote Sensing*, vol. 10, no. 6, p. 840, 2018.
- [71] S. R. Bajracharya, W. Palash, M. S. Shrestha et al., “Systematic evaluation of satellite-based rainfall products over the Brahmaputra basin for hydrological applications,” *Advances in Meteorology*, vol. 2015, p. 1, 2015.
- [72] S. Moazami, S. Golian, Y. Hong, C. Sheng, and M. R. Kavianpour, “Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran,” *Hydrological Sciences Journal*, vol. 61, no. 2, pp. 420–440, 2016.
- [73] X. Tang, J. Zhang, G. Wang et al., “Evaluating suitability of multiple precipitation products for the Lancang river basin,” *Chinese Geographical Science*, vol. 29, no. 1, pp. 37–57, 2019.
- [74] S. Jiang, S. Liu, L. Ren et al., “Hydrologic evaluation of six high resolution satellite precipitation products in capturing extreme precipitation and streamflow over a medium-sized basin in China,” *Water*, vol. 10, no. 1, pp. 25–17, 2017.
- [75] H. Beria, T. Nanda, D. Singh Bisht, and C. Chatterjee, “Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale,” *Hydrology and Earth System Sciences*, vol. 21, no. 12, pp. 6117–6134, 2017.
- [76] M. M. Bitew and M. Gebremichael, “Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model,” *Water Resources Research*, vol. 47, no. 6, pp. 1–11, 2011.
- [77] K. C. Abbaspour, S. A. Vaghefi, and R. Srinivasan, “A guideline for successful calibration and uncertainty analysis for soil and water assessment: a review of papers from the 2016 international SWAT conference,” *Water*, vol. 10, 2017.
- [78] S. Adla, S. Tripathi, and M. Disse, “Can we calibrate a daily time-step hydrological model using monthly time-step discharge data?” *Water*, vol. 11, no. 9, pp. 1750–1825, 2019.
- [79] Q. Zhu, W. Xuan, L. Liu, and Y.-P. Xu, “Evaluation and hydrological application of precipitation estimates derived from PERSIANN-CDR, TRMM 3B42V7, and NCEP-CFSR over humid regions in China,” *Hydrological Processes*, vol. 30, no. 17, pp. 3061–3083, 2016.
- [80] T. Vu, L. Li, and K. Jun, “Evaluation of multi-satellite precipitation products for streamflow simulations: a case study for the Han River Basin in the Korean Peninsula, East Asia,” *Water*, vol. 10, no. 5, p. 642, 2018.
- [81] E. Nkiaka, N. R. Nawaz, and J. C. Lovett, “Effect of single and multi-site calibration techniques on hydrological model performance, parameter estimation and predictive uncertainty: a case study in the Logone catchment, Lake Chad basin,” *Stochastic Environmental Research and Risk Assessment*, vol. 32, no. 6, pp. 1665–1682, 2018.
- [82] N. K. Shrestha, F. M. Qamer, D. Pedreros, M. S. R. Murthy, S. M. Wahid, and M. Shrestha, “Evaluating the accuracy of Climate Hazard Group (CHG) satellite rainfall estimates for precipitation based drought monitoring in Koshi basin, Nepal,” *Journal of Hydrology: Regional Studies*, vol. 13, pp. 138–151, 2017.
- [83] Z. Li, D. Yang, and Y. Hong, “Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River,” *Journal of Hydrology*, vol. 500, pp. 157–169, 2013.
- [84] D. N. Moriasi, J. G. Arnold, M. W. V. Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations,” *Transactions of the ASABE*, vol. 50, pp. 885–900, 2007.

- [85] S. Prakash, A. K. Mitra, I. M. Momin et al., "Seasonal intercomparison of observational rainfall datasets over India during the southwest monsoon season," *International Journal of Climatology*, vol. 35, no. 9, pp. 2326–2338, 2015.
- [86] H. L. Shah and V. Mishra, "Uncertainty and bias in satellite-based precipitation estimates over Indian subcontinental basins: implications for real-time streamflow simulation and flood prediction," *Journal of Hydrometeorology*, vol. 17, no. 2, pp. 615–636, 2016.
- [87] T. Venkata Lakshmi Kumar, H. Alves Barbosa, M. Kumar Thakur, and F. Paredes-Trejo, "Validation of satellite (TMPA and IMERG) rainfall products with the IMD gridded data sets over monsoon core region of India," *Satellite Information Classification and Interpretation*, pp. 1–12, 2019.
- [88] W. Abera, L. Brocca, and R. Rigon, "Comparative evaluation of different satellite rainfall estimation products and bias correction in the Upper Blue Nile (UBN) basin," *Atmospheric Research*, vol. 178–179, pp. 471–483, 2016.
- [89] V. Thiemi, R. Rojas, M. Zambrano-Bigiarini, V. Levizzani, and A. De Roo, "Validation of satellite-based precipitation products over sparsely gauged African River basins," *Journal of Hydrometeorology*, vol. 13, no. 6, pp. 1760–1783, 2012.
- [90] T. Dinku, P. Ceccato, and S. J. Connor, "Challenges of satellite rainfall estimation over mountainous and arid parts of east Africa," *International Journal of Remote Sensing*, vol. 32, no. 21, pp. 5965–5979, 2011.
- [91] T. Dinku, C. Funk, P. Peterson et al., "Validation of the CHIRPS satellite rainfall estimates over eastern Africa," *Quarterly Journal of the Royal Meteorological Society*, vol. 144, no. S1, pp. 292–312, 2018.
- [92] M. P. Young, C. J. R. Williams, J. C. Chiu, R. I. Maidment, and S.-H. Chen, "Investigation of discrepancies in satellite rainfall estimates over Ethiopia," *Journal of Hydrometeorology*, vol. 15, no. 6, pp. 2347–2369, 2014.
- [93] E. Cattani, A. Merino, and V. Levizzani, "Evaluation of monthly satellite-derived precipitation products over East Africa," *Journal of Hydrometeorology*, vol. 17, no. 10, pp. 2555–2573, 2016.
- [94] P. López López, W. W. Immerzeel, E. A. Rodríguez Sandoval, G. Sterk, and J. Schellekens, "Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the Magdalena river basin in Colombia," *Frontiers in Earth Science*, vol. 6, 2018.
- [95] Q. Zeng, Y. Wang, L. Chen, Z. Wang, H. Zhu, and B. Li, "Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013," *Remote Sensing*, vol. 10, no. 3, p. 168, 2018.
- [96] Z. Duan, J. Liu, Y. Tuo, G. Chiogna, and M. Disse, "Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales," *Science of the Total Environment*, vol. 573, pp. 1536–1553, 2016.