Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: http://ees.elsevier.com

Research papers

Adequacy of Satellite-derived Precipitation Estimate for hydrological modeling in Vietnam basins

Manh-Hung Le [a], Venkataraman Lakshmi [a], John Bolten [b], Duong Du Bui [c]

[a] Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA 22904, USA
[b] NASA Goddard Space Flight Center Greenbelt MD 20771, USA
[c] National Center for Water Resources Planning and Investigation, Hanoi 100000, Viet Nam

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief, with the assistance of Yang Hong

Keywords
Satellite-derived Precipitation Estimate
GPM
SWAT
Vietnam

ABSTRACT

This study evaluates eight Satellite-derived Precipitation Estimate (SPE) datasets, which include uncorrected SPE and gauge-corrected SPE products from Tropical Rainfall Measurement Mission Multi-satellite Precipitation Analysis (TMPA), Global Precipitation Measurement (GPM), Climate Hazards group Infrared Precipitation (CHIRP), and Precipitation Estimation form Remotely Sensed Information using Artificial Neural Networks (PERSIANN). These datasets are utilized with six representative river basins, corresponding to six sub-climate zones in Vietnam, during the period 2002–2017. The evaluations were carried out in two parts: 1) inter-comparison of the SPE products with rain gauges, for the six basins; 2) comparison of streamflow simulations, using the Soil and Water Assessment Tool (SWAT) forced by precipitation from rain gauge and SPE products. The GPM Integrated Multi-satellitE Retrievals for GPM (IMERG) Final run version 06B (GPM IMERGF-V6) exhibited the best overall performances among SPE products, in comparison with the rain gauges for the simulation of streamflow. This study is the first of its kind to validate GPM IMERG products in Vietnam, indicating the strong capability of the new IMERG retrieval algorithms. The CHIRP with stations (CHIRPS) dataset demonstrates a relatively low bias, could benefit long-term water resources planning for droughts. In monthly streamflow simulations, the SPE-driven simulations outperformed rain gauge-driven simulations in a larger basin (North West Region), which has low rain-gauge density. The results of this study could be a guide to determine the suitability of different SPE products for hydrological simulations.

1. Introduction

The major uncertainties in hydrological modeling are associated with incorrect precipitation patterns over space (Sangati and Borga, 2009). Several studies indicated that a better representation of the spatial variability in precipitation could improve model performances (Emmanuel et al., 2012; Lobligeois et al., 2014; Zhao et al., 2013). Rain gauge-, radar-, and satellite-based products are popular methods to estimate precipitation across the globe. Rain gauges are the primary approach to obtain precipitation information, as they measure rainfall by directly on the ground and thus do not need transformation into any type of signal, nor need to be corrected (Kidd, 2001). However, rain gauge networks are often sparse, with irregular spatial coverage. In many parts of the world, they are nonexistent (Mondal et al., 2018; Rana et al., 2015). Moreover, it is often challenging to obtain rain gauge data, especially in developing countries and transboundary river basins, for technical and administrative reasons (Gerlak et al., 2011; Plengsaeng et al., 2014).

Ground-based radar systems are useful, and provide data with high temporal and spatial resolution. However, radar systems frequently have a limited spatial range (Michaelides et al., 2009), and are thus most useful for rapid events, typically in urban hydrology (Thorndahl et al., 2017). In addition, radar sensors are often not feasible in developing countries, due to high installment costs and complex maintenance demands. In an effort to cover large areas over long periods, regionally and globally, Satellite-derived Precipitation Estimate (SPE) products emerge as promising approaches to reflect the spatial pattern and temporal variability of rainfall. Several gridded SPE products have been developed over the last few decades, including PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) (Sorooshian et al., 2000); CMORPH (Climate Prediction Center (CPC) MORPHing technique) (Joyce et al., 2004); GSMaP (Global Satellite Mapping of Precipitation) (Ushio et al., 2009); TMPA (TRMM Multi-satellite Precipitation Analysis) (Huffman et al., 2007); and GPM (Global Precipitation Mission) (Hou et al., 2014). Moreover, several promising datasets incorporating gauge, satellite, and re-analysis observations, such as CHIRPS (Climate Hazards group InfraRed Precipitation with Station data) (Funk et al., 2015), and MSWEP (Multi-Source Weighted-Ensemble Precipitation) (Beck et al., 2017), have also been released.

Many studies have shown that gauge-based hydrological models outperform SPE-based models, in terms of streamflow simulation (Duan et al., 2018; Li et al., 2018; Nguyen et al., 2018). However, SPE-driven hydrological simulations exhibit better performance in simulating streamflow than rain gauge-driven hydrological simula-

https://doi.org/10.1016/j.jhydrol.2020.124820
Received 1 November 2019; Received in revised form 19 Febraury 2020; Accepted 9 March 2020
Available online xxx
0022-1694/

tions, for example, in the Luanhe River (Ren et al., 2018), and Lower Mekong River basins (Luo et al., 2019; Mohammed et al., 2018). This is likely associated with the low density of rain gauges and poor-quality of ground rainfall data. For example, a low rain gauge density was observed at the Upper Yangtze River Basin of China, where stations were located approximately every 30,000 km^2 (Liu et al., 2017). Also, Le and Pricope (2017) reported the case of the Nzioa Basin, Western Kenya, where rain gauge data was missing (30–65% records). Interpolating that data resulted in a poorer performance than that of the Climate Forecast System Reanalysis (Saha et al., 2010) and the CHIRPS (Funk et al., 2015) datasets, in terms of streamflow simulation. Wang et al. (2016) indicated that satellite-based rainfall could be more suitable for driving distributed hydrologic models, particularly in basins with poor rain gauge conditions. The superiority of remote sensing in deriving precipitation products has become more pronounced as advanced algorithms have been developed. For example, the TMPA 3B42V7 has proven to be better, compared to its previous version TMPA 3B42V6 (Zhang et al., 2019). The increased spatial and temporal resolution of GPM IMERG follow the successes of TMPA (Hou et al., 2014), with an increase from 0.25° and 3 h to 0.1° and half hour. Furthermore, a fine spatial scale of CHIRPS (0.05°, ~5 km) has been developed (Funk et al., 2015). These developments enable SPE to better characterize the spatial and temporal variability of precipitation.

Each of the SPE products contains various versions, often divided into two groups: gauge-adjusted (gauge-corrected SPE) and gauge-unadjusted (gauge-uncorrected SPE). Gauged-corrected SPE datasets use measured rain gauges or re-analysis data to adjust precipitation estimates at the locations of the gauges. Correction factors used at those rain gauge locations are then applied to the entire dataset, leading to a decrease or increase in rainfall estimates, so that the dataset fits the directly measured, more precise rain gauge data (Beck et al., 2018). However, the gauge networks used for corrections (e.g., GPCC-Global Precipitation Climatology Centre, CPC-Climate Prediction Center) were sparse in many areas, typical in developing countries. Therefore, rigorous comparisons between gauge-corrected SPE products and uncorrected SPE products should be performed, specifically in regions where few gauges are used for creating the adjusted SPE.

In this study, six SPE products were evaluated, including the gauge-corrected SPE products (i.e., GPM IMERGF-V6, TMPA 3B42V7, and CHIRPS V2.0) and the uncorrected SPE products (i.e., GPM IMERGE-V6, TMPA 3B42RT, and CHIRP V2.0), on various climate regions of Vietnam. A hydrological model assessment of the SPE was performed, using the SWAT (Soil Water Assessment Tool) hydrological model. This model has demonstrated strong capabilities in hydrologic assessment throughout Vietnam, in many studies (Ha et al., 2018; Vu et al., 2012; Vu et al., 2017). The primary goal of this study is to obtain insight into the performances between gauge-corrected and uncorrected SPE products in: 1) comparisons to the rain gauge data; and 2) simulations of the monthly stream flow. In this paper, Section 2 introduces the case study. Section 3 presents material and methods. Section 4 presents the results and discussions, and the conclusions are presented in Section 5.

2. Watersheds

In this study, six basins with areas ranging from 684 km^2 to 6042 km^2 were selected (Fig. 1), based on the following criteria. Firstly, headwater basins were selected, to reduce the impact of human activities on the flow regime. Secondly, each basin is located entirely within a single climatological region of Vietnam; it allows to thoroughly examine the performance of SPE datasets over Vietnam. These sub-climatological regions include North West (S1); North East (S2); North Delta (S3); North Central (S4); South Central (S5); and Central Highland (S6). These regions were classified based on the duration of the rainy season; the three heaviest rainfall months; differences in solar radiation; and temperature (Nguyen and Nguyen, 2004), and are widely accepted by the climatological community (Nguyen-Xuan et al., 2016; Phan and Ngo-Duc, 2009; Trinh-Tuan et al., 2019b). Annual precipitation across basins ranges from 1400 to 3800 mm. There is a seasonal variability in precipitation in each basin, with 70–85% total rainfall during May-August (MJJA) or September-December (SOND). For example, in the rainy season, MJJA in the S1 region is highly influenced by the summer monsoon, whereas the S4 region is dominated by the winter monsoon (Nguyen-Le et al., 2015). The average elevations of selected basins are also diverse, ranging from 3 m to more than 2000 m above mean sea level.

3. Data and methods

The chosen approach contains two steps: 1) an inter-comparison of SPE products with in-situ rain gauge data; and 2) an evaluation of a hydrological model for monthly streamflow simulation, driven by rain gauge precipitation measurements and SPE products. Below, we describe the data and methodology used for these two steps.

3.1. Ground hydro-meteorological data

The hydro-meteorological data used in this study were obtained from the Vietnam Meteorological and Hydrological Administration (VMHA) (http://kttvqg.gov.vn/) and National Central for Water Resources Planning and Investigation (NAWAPI) (http://nawapi.gov.vn/). The data were recorded, and have undergone quality control at regional meteorological and hydrological services before the post-processed version was delivered to the VMHA. This process depends on region and data types, which are varied from several days to several months (personal communication).

The daily 2002–2017 runoff data at six hydrological stations were collected corresponding to different climatological regions (Xala (XL) of Ma River; Langson (LS) of Kycung River; Hungthi (HT) of Boi River; Nghiakhanh (NK) of Hieu River; Anchi (AC) of Ve River; and Giangson (GS) of Krong Ana River). The data quality of the streamflow was checked and assured with no gaps, during the given study period. Averaged monthly streamflow at different climate zones in Vietnam exhibits high variability in both time and space (Fig. 2). Examining Fig. 2 as we move from the northern part of Vietnam to the south, that is, from climate zone S1 to S6, we observe that the peak of the monthly runoff shifts from August (zones S1 and S2) to September (zones S3 and S4) to November (zones S5 and S6). We also observe that the AC (Ve River) basin of zone S5 has the largest runoff (by a factor of two as compared to the other river basins).

We collected daily 2000–2017 precipitation data from 31 rain gauge stations across six basins (see Supplementary Data S1). There are several rain gauges in each of these basins, and their number ranges from three to seven. The average missing values across all rain gauges were approximately 1.0%. The long-term mean values were used to substitute for the missing data. The rain gauge data were tested for homogeneity, using the double mass curve to exclude systematic errors over time in the datasets. Annual rainfall at each station was compared with the average annual rainfall of surrounding stations to detect inconsistencies. The results indicated no significant difference between the two curves at all rain gauge stations, ensuring consistency through recorded precipitation.

Besides, at each basin, one to three air temperature datasets (minimum and maximum variables) were collected at meteorological stations, with the same duration as that of precipitation measured by the rain gauges. Since air temperature is less varied than precipitation, a small number of air temperature stations are adequate to represent temperature profiles throughout the basins.

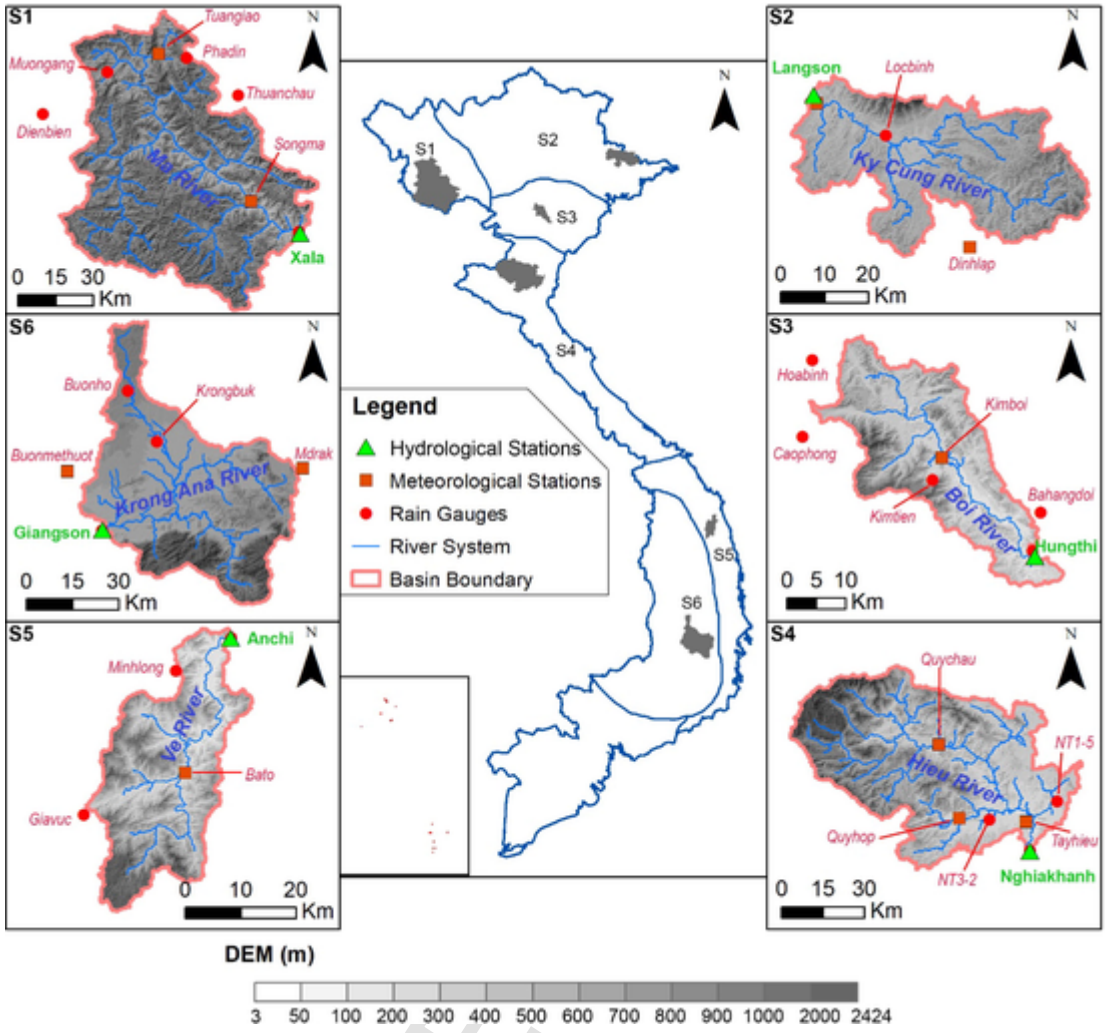

Fig. 1. Digital Elevation Model (DEM) and the distribution of hydrometeorological stations, at six basins, used in this study. S1 North West (XL basin of Ma River); S2 North East (LS basin of Kycung River); S3 North Delta (HT basin of Boi River); S4 North Central (NK basin of Hieu River); S5 South Central (AC basin of Ve River); and S6 Central Highland (GS basin of Krong Ana River).

In conclusion, the data from rain gauges used in this study serve two purposes: 1) as a benchmark to compare with the SPE datasets; and 2) together with air temperature data, as inputs to the SWAT hydrological model for the simulations of streamflow.

3.2. Satellite precipitation Estimation (SPE) products

3.2.1. TMPA precipitation datasets

The Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), launched in late 1997, is a collaboration between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). It is the first space mission to measure rainfall in tropical regions. The TRMM is a low-Earth orbit satellite, equipped with Precipitation Radar (PR); TRMM Microwave Imager (TMI); Visible and Infrared Scanner (VIRS); and Lighting Imaging Sensor (LIS) (Huffman et al., 2007). The two TMPA products used in this study are the near real-time version TMPA 3B42RT (hereafter 3B42RT); and an adjusted version, using monthly gauge precipitation, TMPA 3B42V7 (hereafter 3B42V7) (Huffman and Bolvin, 2013; Huffman et al., 2007). The three hours 0.25° grid TMPA products were accessed from NASA's Goddard Space Flight Center website (https://pmm.nasa.gov/data-access/downloads/trmm), then accumulated to a daily time step.

3.2.2. GPM IMERG precipitation datasets

The Global Precipitation Measurement (GPM) mission was developed as a continuation and improvement of the TRMM mission. The Integrated Multi-satellitE Retrievals for GPM (IMERG) product, is the Level 3 multi-satellite precipitation algorithm of GPM, which combines all of the microwave sensors in the constellation, and Infrared-based observations from geosynchronous satellites (Hou et al., 2014). The two latest products of GPM IMERG used in this study are GPM IMERG Early Run Version 6 (hereafter IMERGE-V6) and IMERG Final Run Version 6 (hereafter IMERGF-V6). The half-hour 0.1° gridded GPM IMERG products were accessed from NASA's Goddard Space Flight Center website (https://pmm.nasa.gov/data-access/downloads/gpm), then accumulated to a daily time step.

3.2.3. CHIRPS precipitation datasets

University of California-Santa Barbara's Climate Hazards Group developed the Climate Hazards group Infrared Precipitation (CHIRP), and the Climate Hazards group Infrared Precipitation with Stations (CHIRPS) datasets, which each provides a more than 30 years quasi-global rainfall dataset. These products aim to support the United States Agency for International Development Famine Early Warning System Network (FEWS NET). The CHIRP dataset estimates rainfall from in-

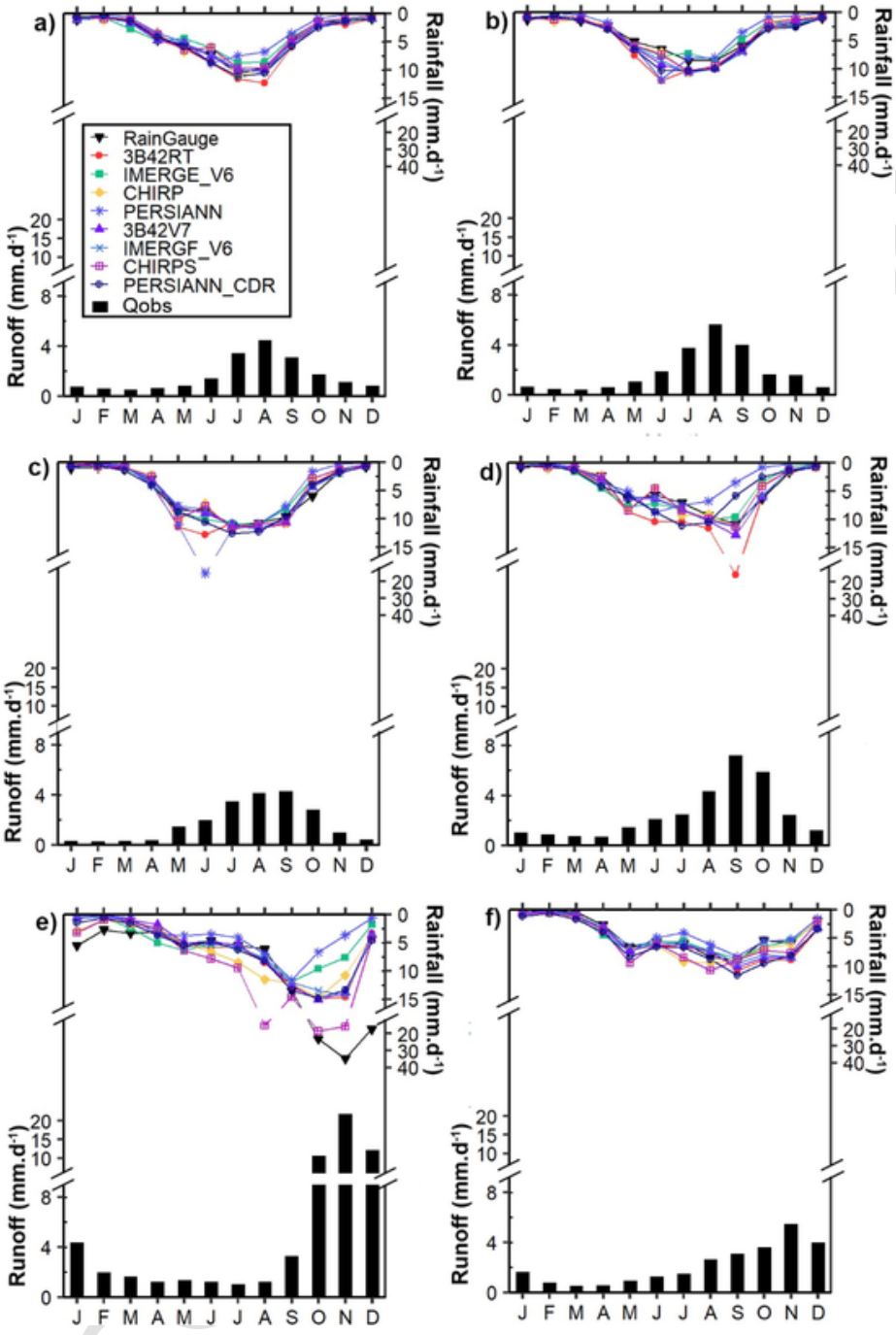

Fig. 2. The observed monthly average runoff and different precipitation datasets (rain gauge; 3B42RT; IMERGE-V6; CHIRP; 3B42V7; IMERGF-V6; and CHIRPS), at the river outlets of a) XL, b) LS, c) HT, d) NK, e) AC, and f) GS basins.

frared cold cloud duration (CCD) regression, calibrated by 2000–2013 TMPA pentadal precipitation product (Funk et al., 2015). The gauge-corrected grid CHIRPS dataset uses rain gauge station observations from various datasets, mainly in the USA, Central America, South America, and sub-Saharan Africa (Funk et al., 2015). This study obtained the daily 0.05° grid CHIRP V2.0 (hereafter CHIRP) and CHIRPS V2.0 (hereafter CHIRPS) datasets from the Climate Hazards Group website (http://chg.geog.ucsb.edu/data/chirps/).

3.2.4. PERSIANN precipitation datasets

Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) is developed at the Center for Hydrometeorology and Remote Sensing (CHRS) at the University of California, Irvine. This product uses artificial neural networks (ANNs) to estimate rainfall rates from cloud-top temperature, measured by long wave infrared imagery at a spatial resolution of 0.25° (Sorooshian et al., 2000). The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) is PERSIANN's adjusted version using Global Precipitation Climatology Project (GPCP) monthly product version 2.2. PERSIANN-CDR has a long-term data set with more than 30 years of data from 1983 to the near present. However, this dataset degrades the temporal resolution to daily scale (Ashouri et al., 2015; Nguyen et al., 2019). This study acquired the daily 0.25° gridded PERSIANN

and PERSIANN-CDR datasets from CHRS portal website (https://chrsdata.eng.uci.edu/).

A summary of SPE is listed in Table 1, and monthly rainfall distributions of SPE at each basin are presented in Fig. 2. The rainfall for the AC (Ve River) basin of Zone S5 is nearly two to three times of the other five river basins. This is reflected in the monthly runoff, which is twice as large in the AC (Ve River) basin, compared to the other five basins.

3.3. SWAT model and setup

SWAT (Soil and Water Assessment Tool) is a physically based, semi-distributed, eco-hydrological model that operates at various time-steps (i.e., daily, monthly, yearly) to simulate the streamflow, sediment, and water quality of large complex river basins (Arnold et al., 1998). In the SWAT model, the smallest spatial unit is the Hydrologic Response Unit (HRU). Runoff is supposed to be predicted separately for each HRU, then routed to estimate the runoff for each sub-basin, as well as that of the entire basin. A detailed description of the SWAT model can be found in Neitsch et al. (2011).

Determining HRUs requires data on elevation, land use, and soil properties. The 30 m Shuttle Radar Topographic Mission Digital Elevation Model (SRTM DEM), was used to estimate slope and delineate the basin boundary; it was obtained from United States Geological Survey (USGS) Earth Explorer (https://earthexplorer.usgs.gov/). The basin boundaries delineated by SRTM DEM were validated, using a reference from the Vietnamese national basin database. The average error between areas delineated from the SRTM DEM and those from the database was 3%, indicating reliable basin boundaries from the SRTM DEM. The 30 m spatial resolution land-use map representing the year 2010, was obtained from the land use portal for Lower Mekong Basin, which was maintained by SERVIR-Mekong (https://rlcms-servir.adpc.net/en/landcover/). A soil map developed by the Vietnam National Institute for Soil and Fertilizers, at 1:1,000,000 scale (National Institute for Soils and Fertilizers, 2002), was used in this study, resampled from polygons to a 30 m raster file. To prepare associated information of soil properties required in SWAT, a soil database using soil water characteristics equations, following the work of Saxton and Rawls (2006), was created.

Statistical descriptions of elevation, land use, and soil used for the SWAT input, are described in Supplementary Data S2. Generally, evergreen forests, mixed forests, and orchards dominate land use, while Acrisols (ACf, ACu), Ferralsols (FRr), and Fluvisols (FLd) are the dominant soil types across basins. In this study, the watershed networks, sub-basins, and HRUs were generated by the QSWAT version 1.7 plug-in in Quantum Geographical Information System (QGIS) version 2.6.1 (Dile et al., 2016). Several advantages of these software systems have been observed, compared to the commonly used Arc SWAT plug-in on the ArcGIS software (Mohammed et al., 2018; Tuo et al., 2016). The advantages are that QSWAT and QGIS are open source software, and QSWAT has additional features such as merging small sub-basins and static, and dynamically visualizing the outputs.

A contributing area over a threshold of 25 km^2 was applied for all basins, resulting in ranges from 15 (HT) to 145 (XL) sub-basins. To create HRUs, the method of the filter by land use, soil, and slope was used, with a threshold of 10% percent of sub-basins chosen for each feature (see Supplementary Data S2). Because solar radiation is not well-observed, we used the simple Hargreaves method (Hargreaves and Samani, 1982), which requires only air temperature data to calculate potential evapotranspiration. To simulate surface runoff processes, the SCS curve number (USDA Soil Conservation Service, 1972) and Variable Storage Routing method (Williams, 1969), were used. By changing precipitation input datasets, including rain gauges; 3B42RT; IMERGE-V6; CHIRP; 3B42V7; IMERGF-V6; and CHIRPS for the SWAT model, seven simulation scenarios were established for each basin to investigate the effects of different rainfall inputs on monthly streamflow simulation.

This study ran the SWAT model on both daily and monthly time scale, selecting the first two years (2000–2001) as the warm-up period; the next eight years (2002–2009) as the calibration period; and the last eight years (2010–2017) as the validation period. The calibration procedure was performed separately for each precipitation dataset. The automatic calibration was performed for streamflow simulation based on the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) (Abbaspour et al., 2007), using the SWAT-CUP tool (Abbaspour et al., 2015). Fifteen sensitive parameters were identified and set up for the same initial range for all scenarios (see Supplementary Data S3). For each scenario, a total of 1000 simulations were generated for the calibration process, using the Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe (1970)) as the objective function.

3.4. Performance metrics

To compare SPE datasets and ground observations, we considered the following three performance metrics in terms of rainfall detection: 1) Probability of Detection (POD); 2) False Alarm Ratio (FAR); and 3) Critical Success Index (CSI). To evaluate the SPE datasets in terms of temporal dynamics, we considered the following three performance metrics: 1) Correlation Coefficient (CC); 2) Relative Bias (RB); and 3) Root Mean Square Error (RMSE). To evaluate hydrological model performance, we considered performance metrics, including Nash-Sutcliffe Efficiency (NSE) and Percentage Bias (PBIAS) (Moriasi et al., 2015). The POD provides the ratio of the total precipitation events, which SPE products detect among the actual precipitation events. The FAR evaluates the fraction of false rainfall events, detected by SPE products from the total rainfall events. The CSI, which is a function of POD and FAR, is the most balanced and accurate detection metric. The rainfall day threshold in this study was set as 0.6 $mm.day^{-1}$ (NCHMF, 2000). The CC is a score of the similarity between the SPE products and ground observations, while the RB and RMSE demonstrate the bias and error of satellite estimates. The NSE indicates how well the observed streamflow and the simulated streamflow fits the 1:1 line. The PBIAS measures the average tendency of the simulated streamflow to be larger

Table 1
Summary of Satellite Precipitation Estimation datasets used in this study, with spatial-temporal characteristics and used period.

Product Name	Spatial coverage	Spatial resolution	Temporal coverage	Finest Temporal resolution	Latency	References
TMPA 3B42RT	50°N – 50°S	0.25°	2000 – present	Every three hours	Several hours	Huffman et al., 2007
GPM IMERGE-V6	65°N – 65°S	0.1°	2000 - present	Every half hour	Several hours	Hou et al., 2014
CHIRP V2.0	50°N – 50°S	0.05°	1981 - present	Daily	Several days	Funk et al, 2015
PERSIANN	60°N – 60°S	0.25°	2000 - present	Hourly	Several days	Sorooshian et al, 2000
TMPA 3B42V7	50°N – 50°S	0.25°	1998 – present	Every three hours	Several months	Huffman and Bolvin, 2013
GPM IMERGF-V6	65°N – 65°S	0.1°	2000 - present	Every half hour	Several months	Hou et al., 2014
CHIRPS V2.0	50°N – 50°S	0.05°	1981 - present	Daily	Several days	Funk et al, 2015
PERSIANN CDR	60°N – 60°S	0.25°	1983 - present	Daily	Several months	Ashouri et al, 2015

or smaller than its observed counterpart. The formulae and perfect scores for each performance metric are given in Table 2.

This study aims to assess the quality of SPE in seasonal water balance. Therefore, we used one-way Analysis of Variance (ANOVA) and Dunnett's Test (Ott and Longnecker, 2015), to compare mean values of SPE with that of the rain gauges. The one-way ANOVA test was first applied to determine whether significant differences exist between the means of rainfall datasets. If the difference was significant, Dunnett's test was used. By comparing each SPE dataset with a single control (rain gauge), it was possible to specify which SPE dataset was significantly different from that of the rain gauge.

4. Results and discussion

4.1. Inter-comparison between rain gauges and satellite precipitation estimate (SPE) datasets

To assess the statistical characteristics of SPE, the precipitation data from the eight SPE products (3B42RT, IMERGE-V6, CHIRP, PERSIANN, 3B42V7, IMERGF-V6, CHIRPS, and PERSIANN-CDR) were directly compared to the precipitation data from rain gauges in the six river basins. We matched precipitation values extracted from SPE's grids to the rain gauge locations. If there were more than one rain gauge located in a grid, we averaged values from those gauges before the comparison. As we examined the GPCC gauges in Vietnam, 27/31 (~90%) rain gauges in our study were not used in the generation of the GPCC product. Therefore, our comparisons between SPE datasets and rain gauges are considered as an independent evaluation.

4.1.1. Detection metrics assessment

Regarding rainfall detection metrics, the gauge-corrected IMERGF-V6 exhibited the best overall performance for the entire period (median POD of 0.718-rank 2; median FAR of 0.391-rank 1; median CSI of 0.505-rank 1; see Fig. 3 and Table 3). The second-best dataset for the entire period was IMERGE-V6 (median POD of 0.699-rank 4; median FAR of 0.403-rank 3; median CSI of 0.497-rank 2), reflecting the quality of the new IMERG retrieval algorithms on both real-time and research products (Huffman et al., 2014; Huffman et al., 2018). Note that uncorrected CHIRP obtained the highest POD score (median POD of 0.878), but the poorest FAR score (median FAR of 0.546), reflecting the imbalance of this rainfall retrieval algorithm. All SPE products exhibited better rainfall detection scores in the wet season than the dry season, which is in line with previous studies (Le et al., 2018; Li et al., 2019; Wang and Lu, 2016). The average median values during the dry season for POD, FAR, and CSI were 0.486, 0.555, and 0.264, respectively. The average median values during the wet season for POD, FAR, and CSI were 0.797, 0.388, and 0.509, respectively (see Fig. 3 and Table 3). In conclusion, the newly released IMERG-V6 dataset outperformed other SPE datasets in terms of rainfall detection. The worst SPE performance was observed in the PERSIANN dataset.

4.1.2. Temporal dynamic metrics assessment

In the assessment of temporal dynamic metrics (see Fig. 4 and Table 3), the best overall correlation coefficient for the entire period was demonstrated by IMERGF-V6 (median CC of 0.650), followed by IMERGE-V6 (median CC of 0.605). Other SPE datasets exhibited moderate CC scores, ranging from 0.296 to 0.472. The CHIRPS and its uncorrected CHIRP product exhibited the best overall RB scores for the entire period (median RB of 0.0 for each product), in line with Beck et al. (2018). These were followed by the IMERGE-V6 and IMERGF-V6 datasets, with moderate RB scores. The best overall RMSE score for the entire period was achieved by IMERGF-V6 (median RMSE of 12.0 mm. d^{-1}), followed by IMERGE-V6 and CHIRPS (median RMSE of 12.4 mm. d^{-1} and 13.0 mm. d^{-1}). Regarding seasonal assessment, the difference between the dry and wet seasons, in terms of temporal dynamic metrics, was not significant. The average median values during the dry season for CC and RB were 0.373 and −0.066, respectively. The average median values during the wet season for CC and RB were 0.409 and 0.054, respectively (see Fig. 4 and Table 3). The average median value of RMSE during the dry season was 7.01 mm. d^{-1}, which was lower than the RMSE figure during the wet season (18.03 mm. d^{-1}). This reflects less variability in rainfall during the dry season, compared to the wet season. The AC basin (South Central), exhibited the largest negative RB and extremely high RMSE values of all SPE products. This is attributed to the greater spatial-temporal rainfall variation of this region (Trinh-Tuan et al., 2019a).

Table 2
Performance metrics for precipitation comparison and hydrological model assessment.

	Statistic	Equation	Optimal Value	Performance Evaluation Criteria
Precipitation Performance Metrics	*POD*	$\frac{N_{11}}{N_{11}+N_{01}}$	1	
	FAR	$\frac{N_{10}}{N_{11}+N_{10}}$	0	
	CSI	$\frac{N_{11}}{N_{11}+N_{01}+N_{10}}$	1	
	CC	$\frac{\sum_{i=1}^{N}(RG_i-\overline{RG})(SPE_i-\overline{SPE})}{\sqrt{\sum_{i=1}^{N}(RG_i-\overline{RG})^2\sum_{i=1}^{N}(SPE_i-\overline{SPE})^2}}$	1	
	RB	$\frac{mean(SPE)}{mean(RG)}-1$	0	
	RMSE	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(SPE_i-RG_i)^2}$	0	
Streamflow Performance Metrics	*NSE*	$1-\frac{\sum_{i=1}^{N}(OBS_i-SIM_i)^2}{\sum_{i=1}^{N}(OBS_i-\overline{OBS})^2}$	1	VG: $NSE \geq 0.80$G: $0.70 \leq NSE<0.80$S: $0.50 \leq NSE<0.70$ NS: $NSE<0.50$
	PBIAS	$1-\frac{\sum_{i=1}^{N}(OBS_i-SIM_i)}{\sum_{i=1}^{N}OBS_i}$	0	VG: $PBIAS \leq \pm 5$ G: $\pm 5<PBIAS \leq \pm 10$ S: $\pm 10<PBIAS \leq \pm 15$ NS: $PBIAS> \pm 15$

Note: N_{11} represents the precipitation observed by the rain gauge and satellite simultaneously. N_{10} represents the precipitation observed by the satellite, but not observed by the rain gauge. N_{01} is contrary to N_{10}. OBS_i is observed streamflow (m^3/s) at the i^{th} day or month, SIM_i is simulated streamflow (m^3/s) at the i^{th} day or month. $\overline{OBS}$ and $\overline{SIM}$ are average observed streamflow and average simulated streamflow, respectively. "VG" Very Good, "G" Good, "S" Satisfactory, "NS" Not Satisfactory.

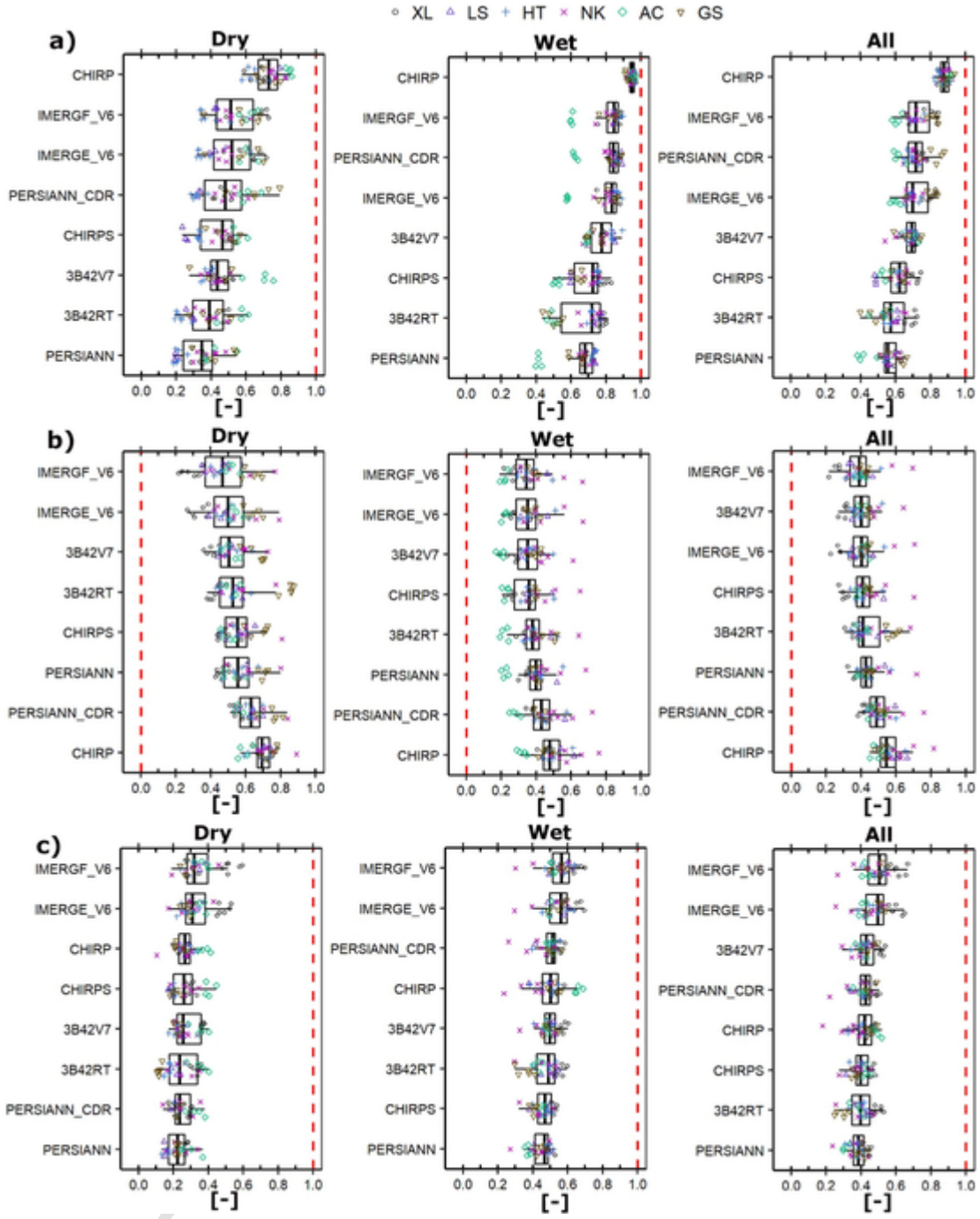

Fig. 3. Box plot of rainfall performance metrics a) POD, b) FAR, and c) CSI for six river basins. The red dash line indicates the optimal value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1.3. Rain-no rain detection assessment

Rain-no rain detection is an important aspect to assess the quality of SPE products. The 3B42V7 and 3B42RT exhibited the most similar figures in terms of the number of rainy days, compared to the observations from rain gauges, for the entire period (Fig. 5). Specifically, 32% of observations across six basins had rainfall events over the entire period; those figures from 3B42V7 and 3B42RT comprised 33% of the entire period. The rainy days detected by PERSIANN, CHIRPS, IMERGF-V6, IMERGE-V6, accounted for 30%, 27%, 36%, and 37% of the entire period, respectively. CHIRP exhibited a large overestimation of rainy days, as 62% of its entire period measured rainfall events, respectively reflecting the impact of intercept values in its algorithm (Funk et al., 2015). The gauge-corrected PERSIANN-CDR also highly overestimated the rainy days as these days accounted for 42% of its entire period. During the dry season, apart from CHIRP, all SPE datasets underestimated rainy days, reflecting the difficulty of SPE in terms of detecting short-term rainfall events. However, IMERG products exhibited significant improvement over the other SPE products. During 19% of the dry period, rain gauge data observed rainfall events, while those estimations from IMERGF-V6 and IMEGRE-V6 were only 2% less (17% of the dry period), suggesting frequently temporal rainfall sampling (every 30 min) could benefit in capturing short-term rainfall events. On the other hand, the IMERG retrieval algorithm highly overestimated rain-

Table 3
Median values of the performance metrics of six Satellite-derived Precipitation Estimation, based on daily rain gauge, during 2002–2017. For all the metrics, except for FAR and RMSE, larger values represent the better performance of SPE products. Values in bold represent the best score for each metric.

		3B42RT	IMERGE_V6	CHIRP	PERSIANN	3B42V7	IMERGF_V6	CHIRPS	PERSIANN_CDR
POD	Dry	0.391	0.519	**0.732**	0.347	0.437	0.515	0.465	0.481
	Wet	0.722	0.835	**0.949**	0.680	0.778	0.846	0.725	0.843
	All	0.572	0.699	**0.878**	0.558	0.693	0.718	0.624	0.717
FAR	Dry	0.527	0.500	0.694	0.558	0.507	**0.467**	0.554	0.633
	Wet	0.380	0.352	0.480	0.400	0.354	**0.344**	0.362	0.431
	All	0.411	0.403	0.546	0.427	0.400	**0.391**	0.410	0.490
CSI	Dry	0.237	0.311	0.267	0.223	0.255	**0.322**	0.261	0.237
	Wet	0.491	0.563	0.503	0.468	0.498	**0.565**	0.470	0.513
	All	0.399	0.497	0.423	0.386	0.430	**0.505**	0.398	0.427
CC	Dry	0.315	0.497	0.302	0.245	0.430	**0.550**	0.341	0.302
	Wet	0.317	0.610	0.316	0.257	0.420	**0.635**	0.364	0.351
	All	0.358	0.605	0.362	0.296	0.472	**0.651**	0.394	0.389
RB	Dry	**−0.02**	0.06	−0.09	−0.41	−0.08	−0.04	**−0.02**	0.07
	Wet	0.18	−0.04	−0.01	0.01	0.07	0.05	**0.00**	0.17
	All	0.17	−0.03	**0.00**	−0.03	0.07	0.05	**0.00**	0.20
RMSE	Dry	9.20	6.60	6.30	7.30	6.80	**6.10**	6.30	7.50
	Wet	21.6	**15.4**	17.2	18.7	20.1	16.2	16.7	18.3
	All	16.5	12.4	13.2	15.0	15.4	**12.0**	13.0	14.4

fall events during the wet season, suggesting a re-evaluation for this algorithm during this period.

4.1.4. Mean annual rainfall assessment

Fig. 6 compares average 2002–2017 rainfall values annually, during the wet season, and the dry season, between rain gauge and SPE products. It was found that CHIRPS product exhibited the most statistically equal mean with rain gauges, among SPE products (78.6% agreeing cases during the entire period). The following SPE products, which demonstrated significantly similar means with those from rain gauges, were CHIRP, IMERGF-V6, IMERGE-V6 (overall agreement 70.2%, 67.5%, and 59.5%, respectively). This finding is in line with the low RB scores of CHIRPS and its uncorrected counterpart, previously described, reflecting that CHIRPS' retrieval algorithm is suitable for trend analysis and drought assessment. From another perspective, all SPE products achieved better agreement during the dry season, than the wet season. The worst mean estimation was observed at the PERSIANN dataset.

4.2. Hydrological simulation driven by different precipitation data inputs

4.2.1. Daily simulations

Fig. 7 represents the performance measures for daily streamflow SWAT simulation, driven by different precipitation datasets. For the NSE scores (Fig. 7a), rain gauge-driven simulations exhibited the best overall performance (median NSE of 0.720). Two SPE-based models had moderate performances in simulating daily streamflow. These are IMERGF-V6-driven simulations; IMERGE-V6-driven simulations, with median NSE scores of 0.600 and 0.500, respectively. Overall, based on the median NSE scores, the rain gauge-based models exhibited *Good* performances in a daily simulation; two SPE-based models (IMERF-V6 and IMERE-V6) demonstrated *Satisfactory* performances; while other SPEs driven simulations performed at *Unsatisfactory* levels (Moriasi et al., 2015).

The daily SPE-driven simulations performed better in terms of the PBIAS score (Fig. 7b). The median PBIAS of IMERG-F-V6-driven simulations was −1.95%, followed by CHIRPS-driven simulations (2.05%); 3B42RT-driven simulations (2.50%); and rain gauge-driven simulations (4.65%). These performances were at *Very Good* levels (Moriasi et al., 2015). The daily simulation using rainfall inputs from PERSIANN-CDR exhibited at *Good* levels based on the PBIAS score (median PBIAS of −6.85%); the PBIAS scores of CHIRP-, and 3B42V7- driven simulations were at *Satisfactory* levels (median PBIAS of 10.10 and 10.25%, respectively). With the median PBIAS scores greater than 15%, IMERGE-V6-, and PERSIANN-driven simulations were at *Unsatisfactory* levels (Moriasi et al., 2015). Details of the daily simulation results can be found at Supplementary Data S4.

4.2.2. Monthly simulations

Fig. 8 presents the performance measures for monthly streamflow SWAT simulation, imposed with different precipitation datasets. Regarding the NSE scores (Fig. 8a), the best overall performance was gained by rain gauge-driven simulations (median NSE of 0.875). The gauge-corrected SPE-based models had comparable performances in simulating monthly streamflow. The median NSE values of IMERGF-V6-driven simulations; 3B42V7-driven simulations; CHIRPS-driven simulations, PERSIANN-CDR-driven simulations were 0.770, 0.740, 0.705, and 0.645, respectively. Apart from PERSIANN-driven simulations, the uncorrected SPE-based model produced monthly streamflow at moderate levels. The median NSE values of IMERGE-V6-driven simulations, 3B42RT-driven simulations; and CHIRP-driven simulations were 0.680, 0.545, and 0.540, respectively. Overall, based on the median NSE scores, the rain gauge-based models exhibited *Very Good* performances; the gauge-corrected SPE-based models demonstrated *Good* (IMERGF-V6, CHRIP, 3B42V7) and *Satisfactory* (PERSIANN-CDR) performances; and uncorrected SPE-based models performed at *Satisfactory* (IMERGE-V6, 3B42RT, CHIRP) and *Unsatisfactory* (PERSIANN) levels (Moriasi et al., 2015).

In terms of the PBIAS score (Fig. 8b), the median PBIAS of rain gauge-driven simulations was 1.25%, followed by 3B42V7-driven simulations (1.55%); CHIRPS-driven simulations (1.70%); and IMERGF-V6-driven simulations (2.60%). These performances were at *Very Good* levels (Moriasi et al., 2015). The models using rainfall inputs from uncorrected 3B42RT and CHIRP datasets exhibited at *Good* levels (median PBIAS of −6.25% for 3B42RT and 8.10% for CHIRP); the PBIAS scores of IMERGE-V6 driven simulations were at *Satisfactory* levels (mean PBIAS of 11.2%). The median PBIAS of PERSIANN-driven simulations (24.95%) indicated that these simulations performed at *Unsatisfactory* level. Details of the monthly simulation results can be found at Supplementary Data S5.

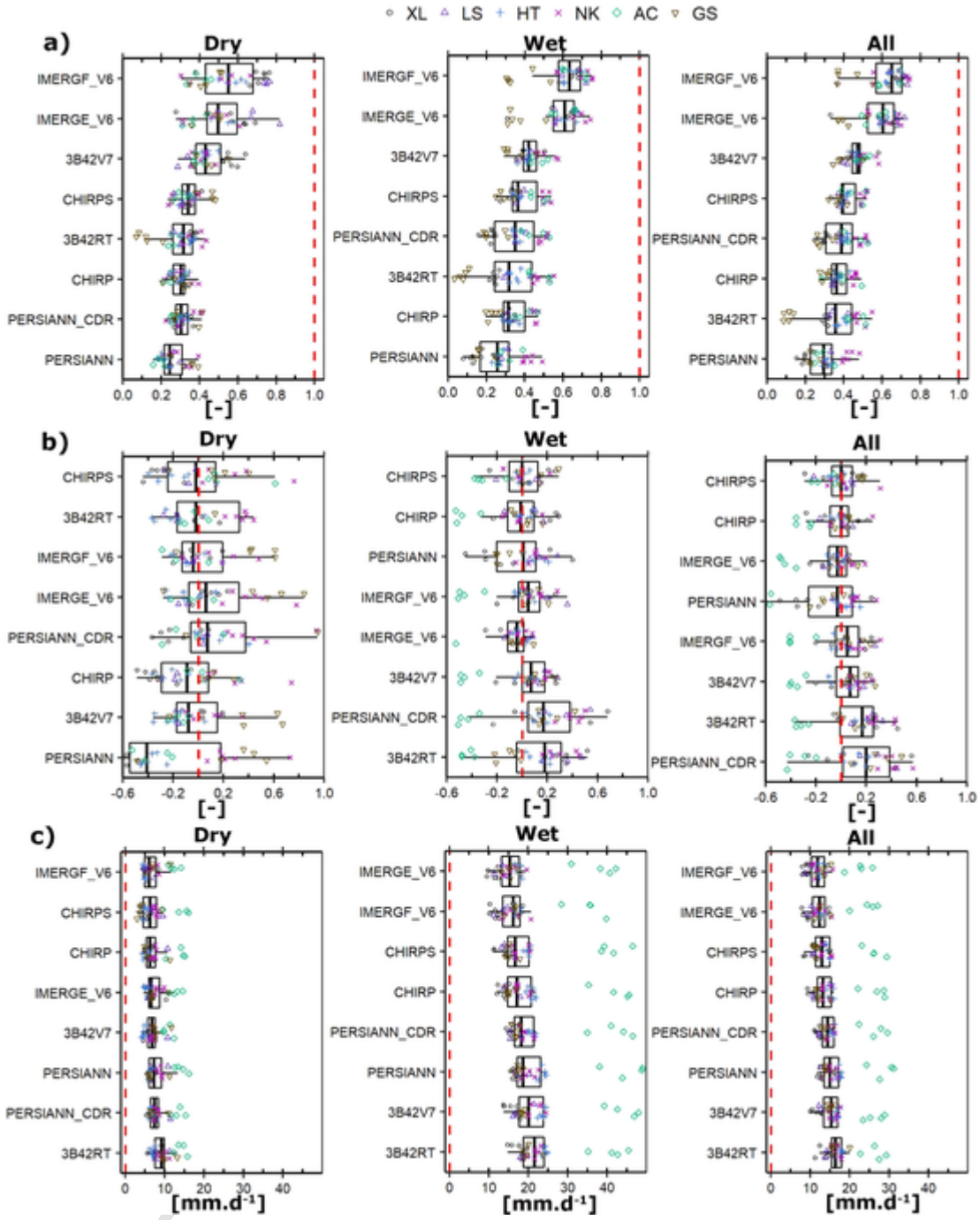

Fig. 4. Box plot of rainfall performance metrics a) CC, b) RB, and c) RMSE for six river basins. The red dash line indicates the optimal value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2.3. SPE-driven simulations in a large basin

Although rain gauge-driven simulations exhibited the best overall performance, compared to SPE datasets, the number of cases in which the PBIAS scores at *Unsatisfactory* level (|PBIAS| greater than 15%) from rain-gauge-driven simulations were high. These *Unsatisfactory* PBIAS scores were found at five and three cases in the daily time step and monthly time step, respectively. This reflects an insufficient estimation at the spatial scale from the rain gauge. On the other hand, the PBIAS's *Unsatisfactory* figures for the IMERGF-V6-based model were observed at two simulations (daily time step, Fig. 7b.) and one simulation (monthly time step, Fig. 8b) only. In daily streamflow simulations at the large XL basin, gauge-corrected SPE-driven simulations exhibited comparable in performance as rain gauge-driven simulations. The daily NSE scores of rain gauge-driven simulations during the calibration and validation period, at the XL basin, were 0.64 and 0.69, respectively (Fig. 9a.). Those figures from IMERGF-V6 were nearly similar, with the scores of 0.63 and 0.69, respectively (Fig. 9c). Interestingly, in monthly streamflow simulation at the XL basin, the SPE-based models were even slightly better than rain-gauge driven simulation. The daily NSE scores of rain gauge-driven simulations during the calibration and validation period, at the XL basin, were 0.80 and 0.74, respectively (Fig. 10a); those from the IMERGF-V6-driven simulations were 0.84 and 0.91, respectively (Fig. 10c). We also examined the ex-

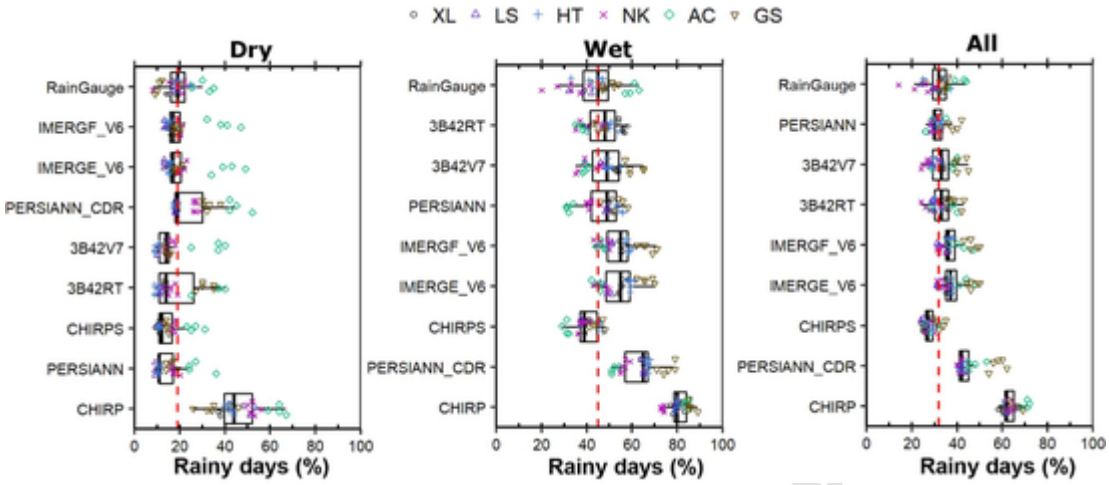

Fig. 5. Box plot of the number of rainy days retrieved from rain gauge and Satellite-derived Precipitation Estimation during the dry, the wet, and entire period. The red dash line indicates the median value from the rain gauge. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

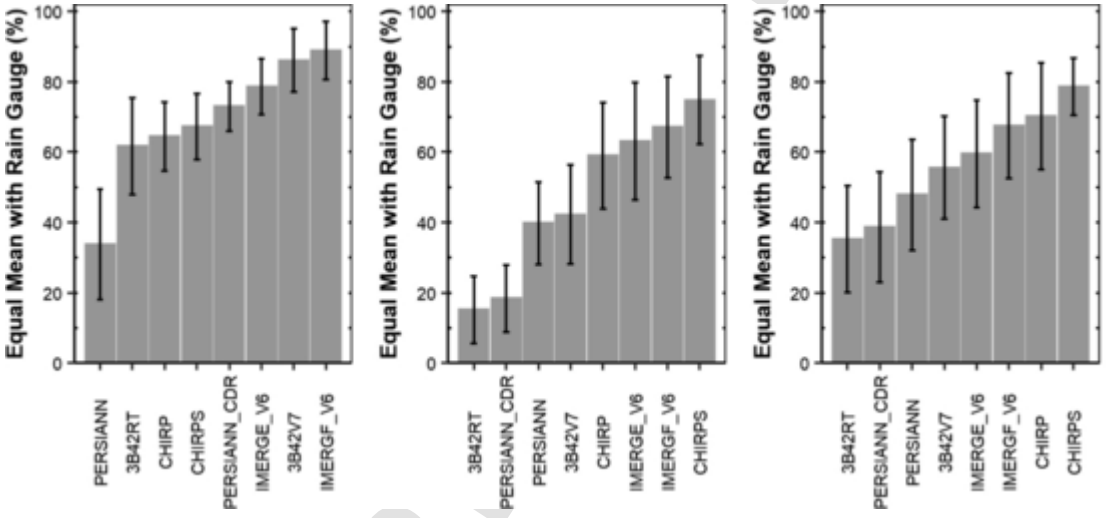

Fig. 6. Statistically equal mean between Satellite-derived Precipitation Estimation products and rain gauge, during the dry, the wet, and the entire period (2002–2017).

ceedance probability of both daily and monthly streamflow, from observations and simulations driven by different precipitation datasets, at the XL basin (Figs. 11 and 12). Overall, the flow curves from simulated results followed the observation curves well, at low exceedance levels. At high exceedance level flow, the simulated curves began to look different from the observed curve. CHIRPS- and CHIRP-driven simulations produced better accurate curves than that from rain gauge-driven simulation, up to exceedance of around 85% flow for daily streamflow data and around 75% flow for monthly streamflow data, suggesting the capability of those products in terms of low-flow simulation.

4.2.4. SPE-driven simulations in basin frequently affected by typhoon and tropical storm

Many poor scores were reported for the AC (Ve river) basin when we used SPE datasets as inputs to the SWAT model; whereas the rain gauge-driven simulations exhibited *Good* to *Very Good* performances in daily streamflow simulation and monthly streamflow simulation (daily NSE for validation: 0.72, monthly NSE for validation: 0.88; see Supplementary Data S4-5). We initially expected that when the SWAT model was recalibrated with satellite precipitation data, the problem of underestimation would be mitigated. However, since the underestimations of the SPE products were extremely large at this basin, the re-calibrated SWAT model did not perform well. This large underestimation can be seen in Fig. 13. We used violin plots to examine the distribution of monthly basin rainfall; monthly streamflow from SPE-driven simulations without re-calibration (i.e., using rain gauge calibration parameters); and monthly streamflow from SPE-driven simulations, with re-calibration from each SPE dataset. Although SPE products demonstrated similar distribution at low to medium rainfall (< 500 mm. month^{-1}), a large discrepancy was found with high rainfall. The maximum rainfall per month, measured by the rain gauge, was up to 2250 mm. month^{-1}; while the figures from SPE datasets, ranged from only 800 to 1500 mm. month^{-1}. When we used the rain gauge's calibrated parameters for models using SPE rainfall inputs, the distributions of simulated streamflow were significantly different from that of observed streamflow (Fig. 13b). On the other hand, by using re-calibrated parameters in each SPE dataset, the distributions of simulated streamflow were more similar to observed streamflow. However, a large dissimilarity between high streamflow distribution from models and observed data has been identified (Fig. 13c). In short, from the simulation results at the AC basin, we suggest a re-evaluation for SPE datasets at regions that are heavily influenced by the tropical cyclone and monsoon systems. The simulated results of SPE-based models without re-calibrated parameters were even worse, compared to the models using re-calibration parameters, which is in line with previous studies (Alazzy et al., 2017; Li et al., 2018).

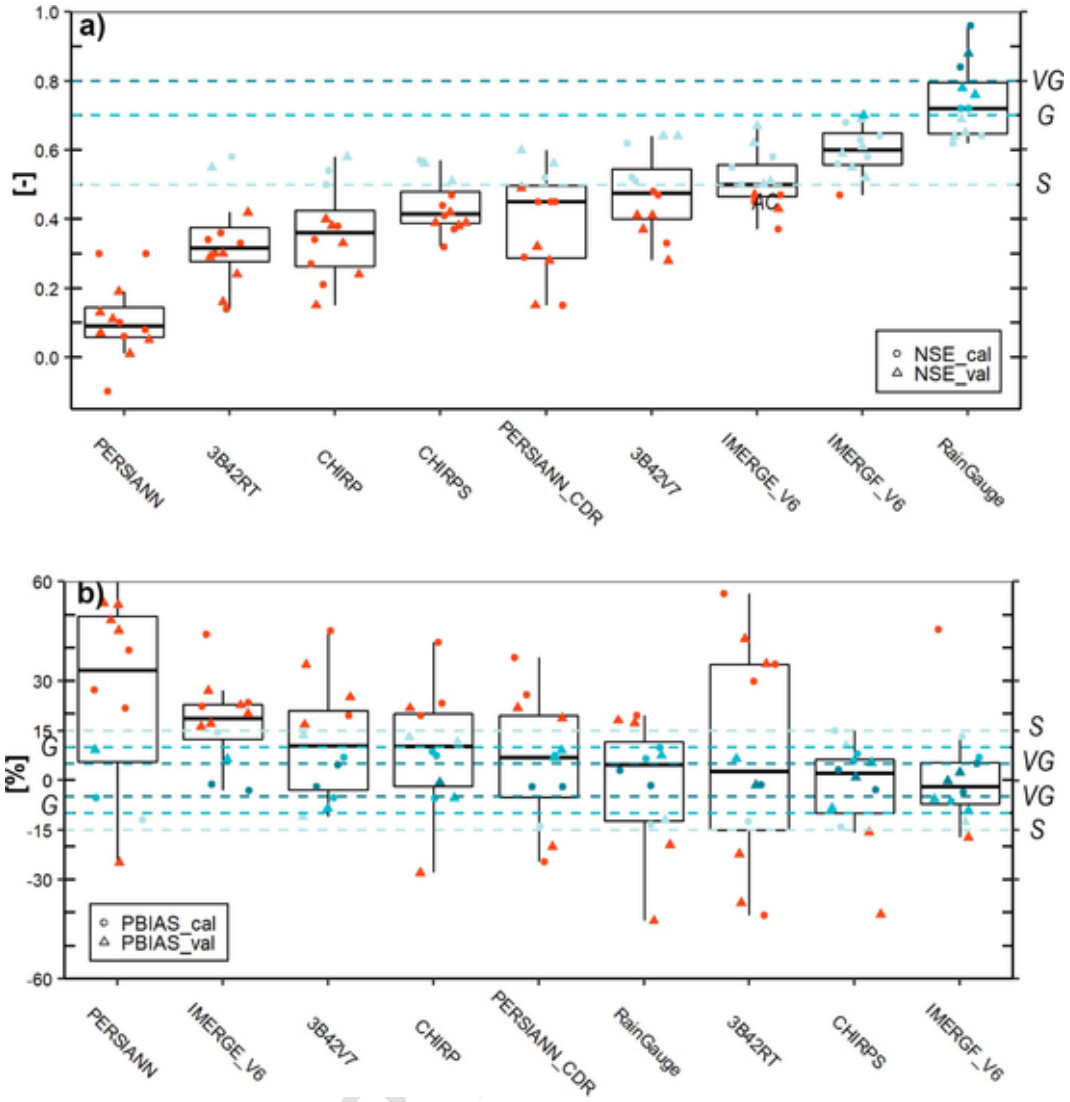

Fig. 7. Performance measures a) NSE, b) PBIAS of daily streamflow SWAT simulations, driven by different precipitation input datasets, at the six basins in Vietnam. Total samples in each boxplot are 12 (six calibration and six validation values). Boxes represent the interquartile range and median and outliers are lower or higher than the 10th or 90th percentile, respectively. The performance explanation: VG Very Good, G Good, S Satisfactory. The evaluation period: Cal. Calibration (2002–2009), Val. Validation (2010–2017).

4.3. Gauge-corrected and uncorrected SPE products

Table 4 presents the differences in the median between gauge-corrected and uncorrected versions of SPE datasets, in terms of precipitation and streamflow performance metrics. The gauge-corrected products incorporated five-day gauge data (CHIRPS) and monthly gauge data (3B42V7, IMERGF-V6, and PERSIANN-CDR) datasets. We expected that the late release of gauge-corrected products (often in several months' latency) would result in these products outperforming the uncorrected products. However, by using various precipitation metrics, we detected that the gauge-corrected products exhibited little improvement, or even worse performances (e.g., CHIRPS – CHIRP for POD: −0.254, PERSIANN-CDR-PERSIAN for RB: +0.230), in a daily time step. This suggests the necessity of incorporating daily gauge observations to improve precipitation performance at this time step. The monthly streamflow performance metrics indicated a considerable improvement on the NSE scores in both daily and monthly simulation (averaged + 0.13 for daily simulation and + 0.20 for monthly simulation), and a significant reduction in the PBIAS scores (-10.3% for daily simulation and −8.6% for monthly simulation). This reflects that corrections provide more benefits to hydrological applications.

Note that, in principle, higher spatial resolution is better. However, CHIRPS uses only infrared data, and, typically, this dataset did not capture well the variability in precipitation in space. Therefore, the theoretical higher spatial resolution might not provide any practical benefit. Previous studies comparing TMPA and CHIRPS performance (Luo et al., 2019; Wu et al., 2018) reached largely similar conclusions. PERSIANN precipitation datasets seem to not be comparable with other SPE products used in this study, probably because these datasets (1) also use infrared data as the CHIRPS dataset; (2) have a relative coarse spatial resolution.

4.4. The relative performance of SPE to rain gauge for SWAT simulation

It is worth highlighting the rationale of "relative performance" when considering the gauge-based model as a benchmark to investigate the adequacy of SPE in driving hydrological modeling. We found evidence that the performances of SPE-based models relative to the rain gauge-based models ($P_{relative}$ = (NSE_{SPE}-NSE_{RG})/NSE_{RG}x100), were to some degree functions of elevation range and rain gauge network den-

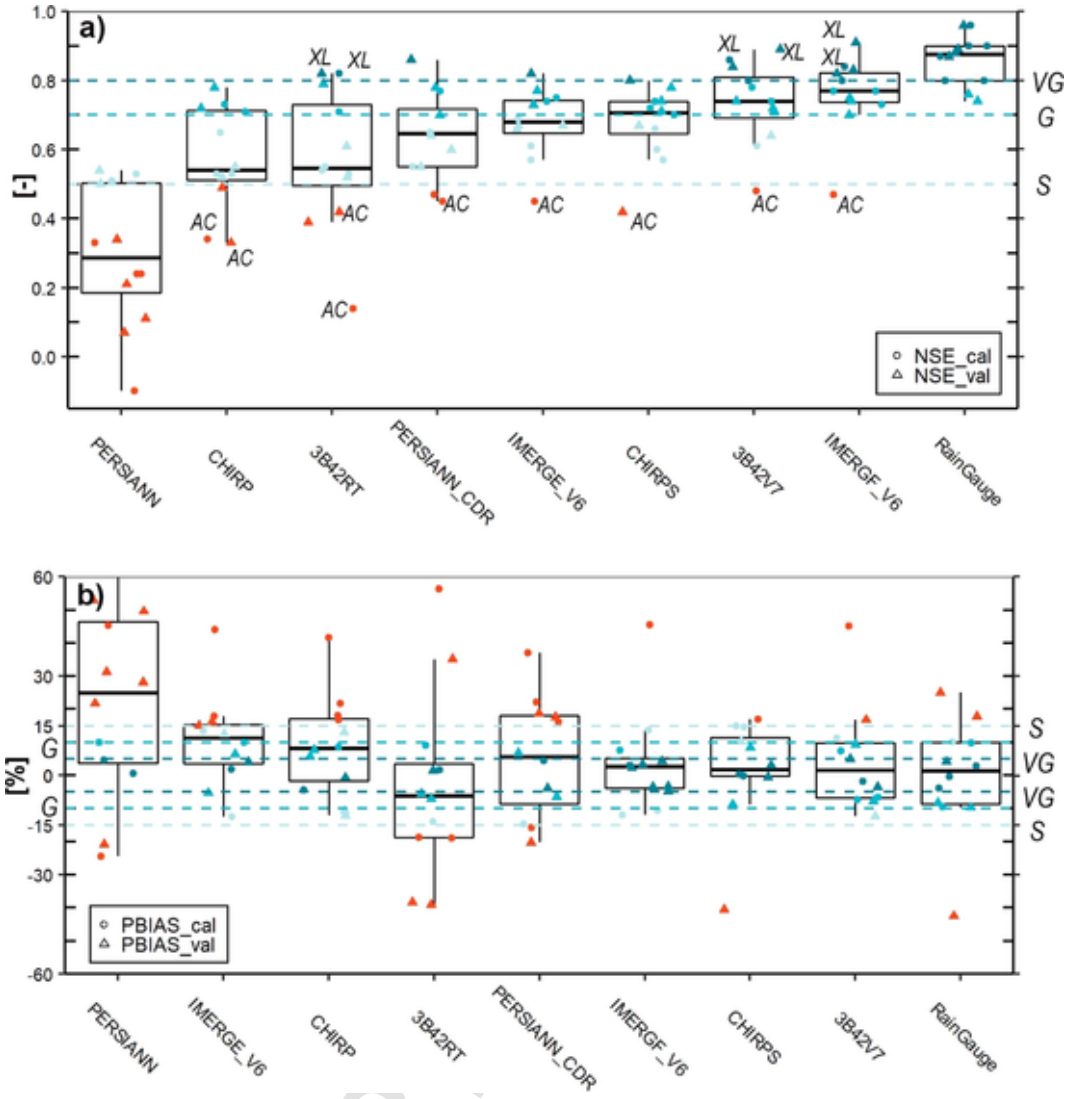

Fig. 8. Performance measures a) NSE, b) PBIAS of monthly streamflow SWAT simulations, driven by different precipitation input datasets, at the six basins in Vietnam. Total samples in each boxplot are 12 (six calibration and six validation values). Boxes represent the interquartile range and median and outliers are lower or higher than the 10th or 90th percentile, respectively. The performance explanation: VG Very Good, G Good, S Satisfactory. The evaluation period: Cal. Calibration (2002–2009), Val. Validation (2010–2017).

sity (Fig. 14). Larger basins tended to be poorly gauged, and streamflow simulations imposed with SPE at large basins had comparable simulation results with those using precipitation from rain gauges. The 3D surface, fitted from elevation range, rain gauge density, and $P_{relative}$, suggested that high $P_{relative}$ values were observed at basins with a low rain-gauge density. On the other hand, several studies (Blöschl, 2013) indicated that hydrological simulations are performed better in large storage of watersheds. Because the relative variation in streamflow at these watersheds is small, it leads to better simulation results, compared to smaller watersheds.

4.5. Limitations and Further study

In this study, we utilized the SPE products for monthly streamflow simulation, using their finest grid; however, we did not use the same grid size for different precipitation inputs. This is due to two factors. Firstly, Bai et al. (2018) revealed that the correlation of SPE with rain gauges, is different from various spatial resolutions. Secondly, the lack of advantage of GPM IMERG on streamflow simulation, compared to 3B42V7 at the Ganjiang River basin, might be due to the resampling of the grid size from 0.1° to 0.25° of GPM IMERG (Zhang et al., 2019).

SPE-driven simulations did not perform well in our simulations at a daily time step. Further study should apply a bias-correction scheme for the SPE products on this time step. The study would greatly benefit examination of extreme analysis and disaster management.

5. Conclusions

This study evaluated the performances of eight Satellite Precipitation Estimation (SPE) datasets, including uncorrected versions (IMERGE-V06, TMPA 3B42RT, CHIRP, and PERSIANN) and gauge-corrected versions (IMERGF-V6, TMPA 3B42V7, CHIRPS, and PERSIANN-CDR), regarding six sub-climate zones of Vietnam. The work consists of two parts: 1) comparisons of the SPE products to rain gauges, and 2) using hydrological SWAT models to simulate monthly streamflow at the six basins, representative of the six climate zones. Our findings can be summarized as follows:

1) The SPE products exhibited a slightly better performance during the wet season, compared to the dry season, in terms of rainfall detection metric (POD, FAR, and CSI). However, the temporal dynamic performance (CC and RB) did not show any significant difference between the two seasons.

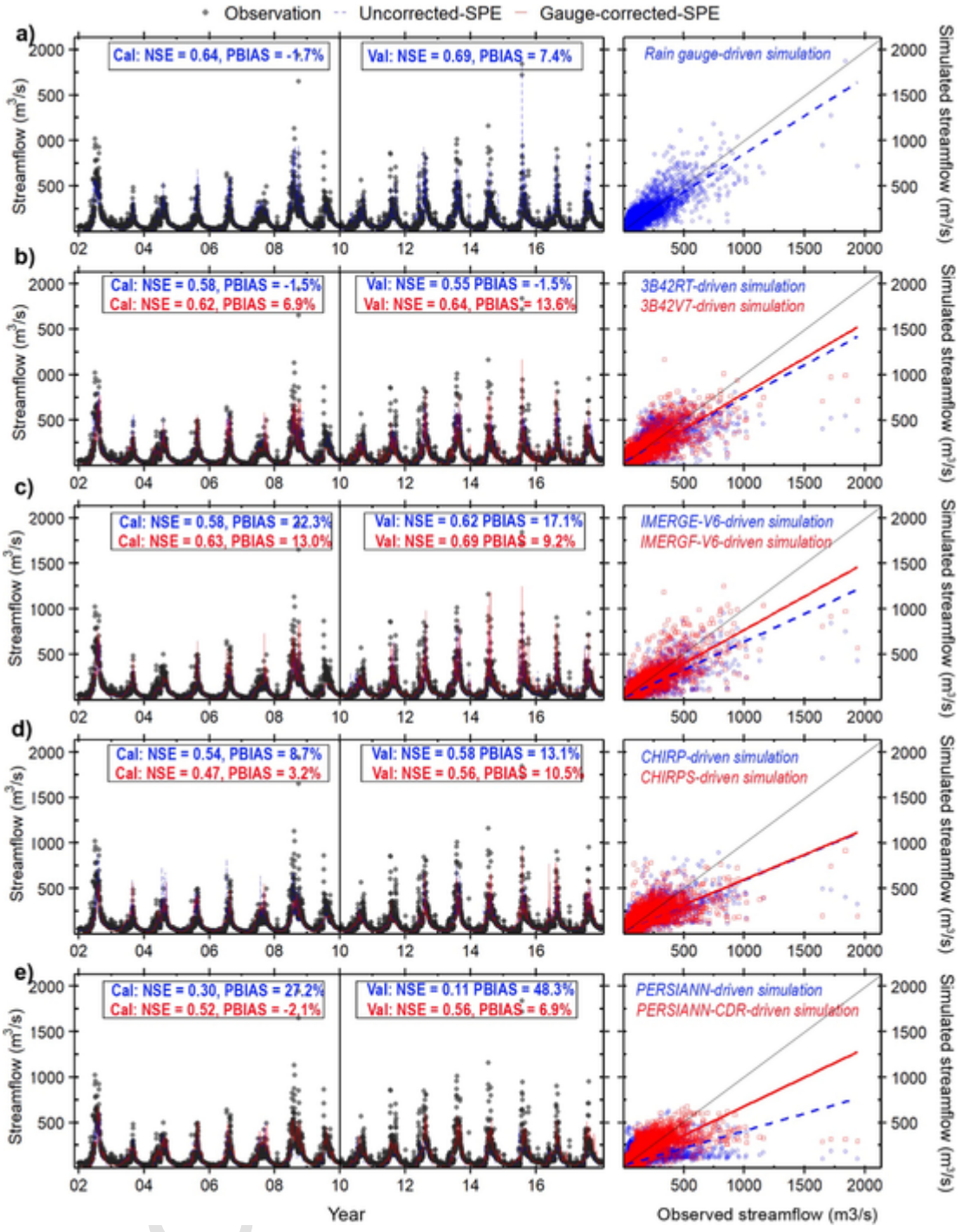

Fig. 9. Comparison between daily observed streamflow and simulated streamflow, driven by a) rain gauge; b) TMPA precipitation datasets; c) GPM IMERG precipitation datasets; d) CHIRPS precipitation datasets; and e) PERSIANN precipitation datasets, at the XL basin, during 2002 – 2017. The calibration period is 2002–2009; the validation period is 2010–2017. Apart from panel a), blue texts denote performances of uncorrected-SPE-driven simulations, while red texts denote performances of gauge-corrected-SPE-driven simulations. In the scatter plot, dash blue line exhibits linear regression between simulated streamflow from uncorrected SPE-based model and observed streamflow. Red line exhibits linear regression between simulated streamflow from gauge-corrected SPE-based model and observed streamflow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2) IMERGF-V6 exhibited the best overall performance among SPE products, in comparison with rain gauges, and as inputs to the SWAT models for streamflow simulations. Our study is the first attempt to evaluate the performance of GPM IMERG in Vietnam, suggesting strong capability for this product in hydrological application purposes.
3) CHIRPS achieved the smallest bias among SPE products, compared to rain gauge data, reflecting the aim of this product as a drought-warning system and for trend analysis.
4) Gauge-corrected versions of SPE products exhibited slightly better over the uncorrected versions of SPE products, in terms of precipitation performance metrics. This suggests that the use of sub-monthly and monthly rain gauges did not significantly benefit SPE's im-

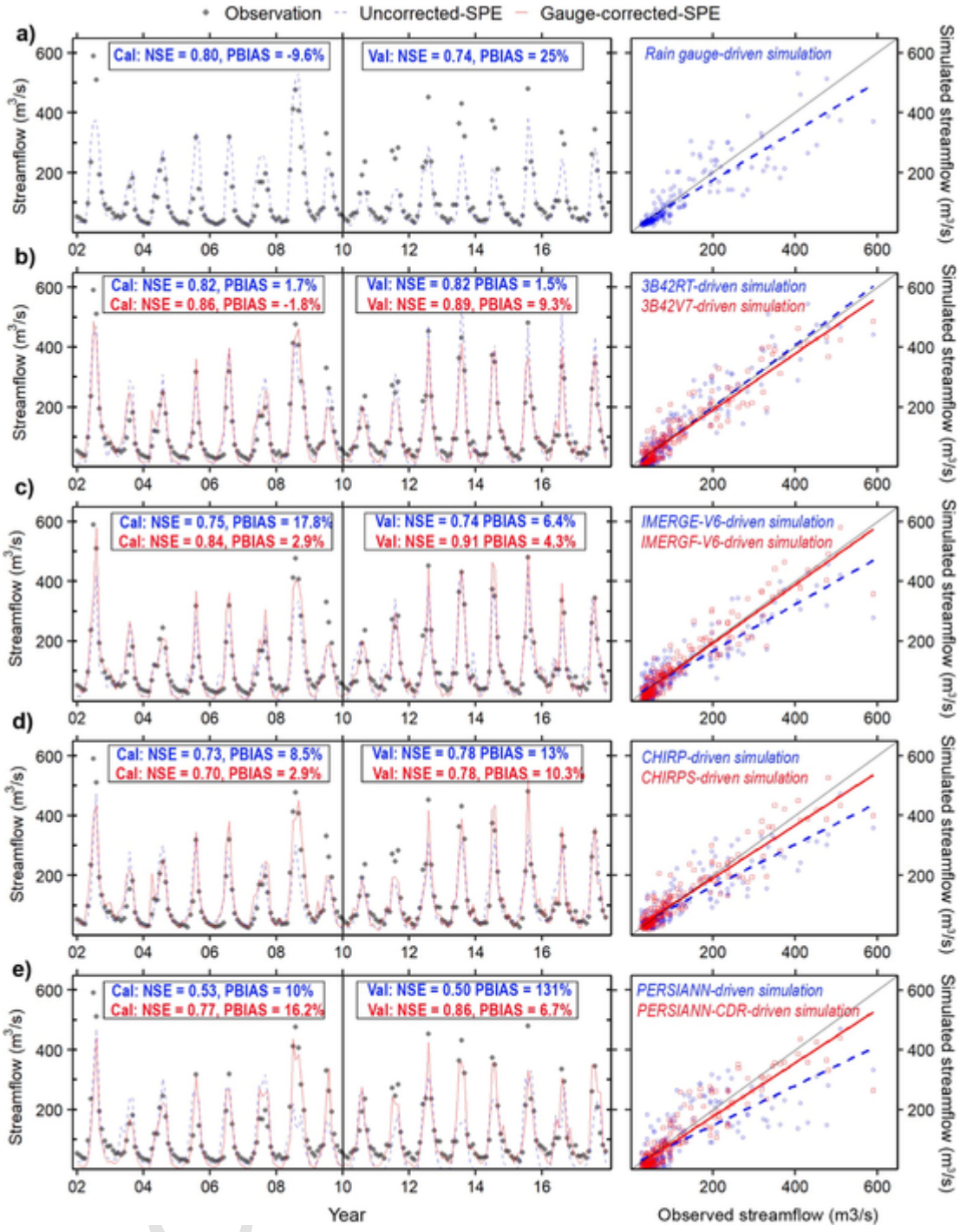

Fig. 10. Similar to Fig. 9 but for monthly simulations.

provement at the daily time step. However, the gauge-corrected SPE products performed better than their uncorrected counterparts in both daily and monthly streamflow simulation.

5) SPE products can serve as alternative inputs to enhance the performance of hydrological models in basins, with a low rain-gauge network density.

This study determines the ability of SPE products to estimate rainfall, and produce input data for streamflow simulations in Vietnam. Our findings could be used as a guide to select which SPE products are suitable for hydrological applications. Although this study is specific for hydro-climatic conditions in the river basins of Vietnam, the methodology can be applied to watersheds in other regions of the world.

CRediT authorship contribution statement

Manh-Hung Le: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Venkataraman Lakshmi:** Supervision, Project administration, Funding acquisition, Writing - review & editing. **John Bolten:** Investigation, Resources, Project administration, Funding acquisition. **Duong Du Bui:** Data curation, Writing - review & editing, Funding acquisition.

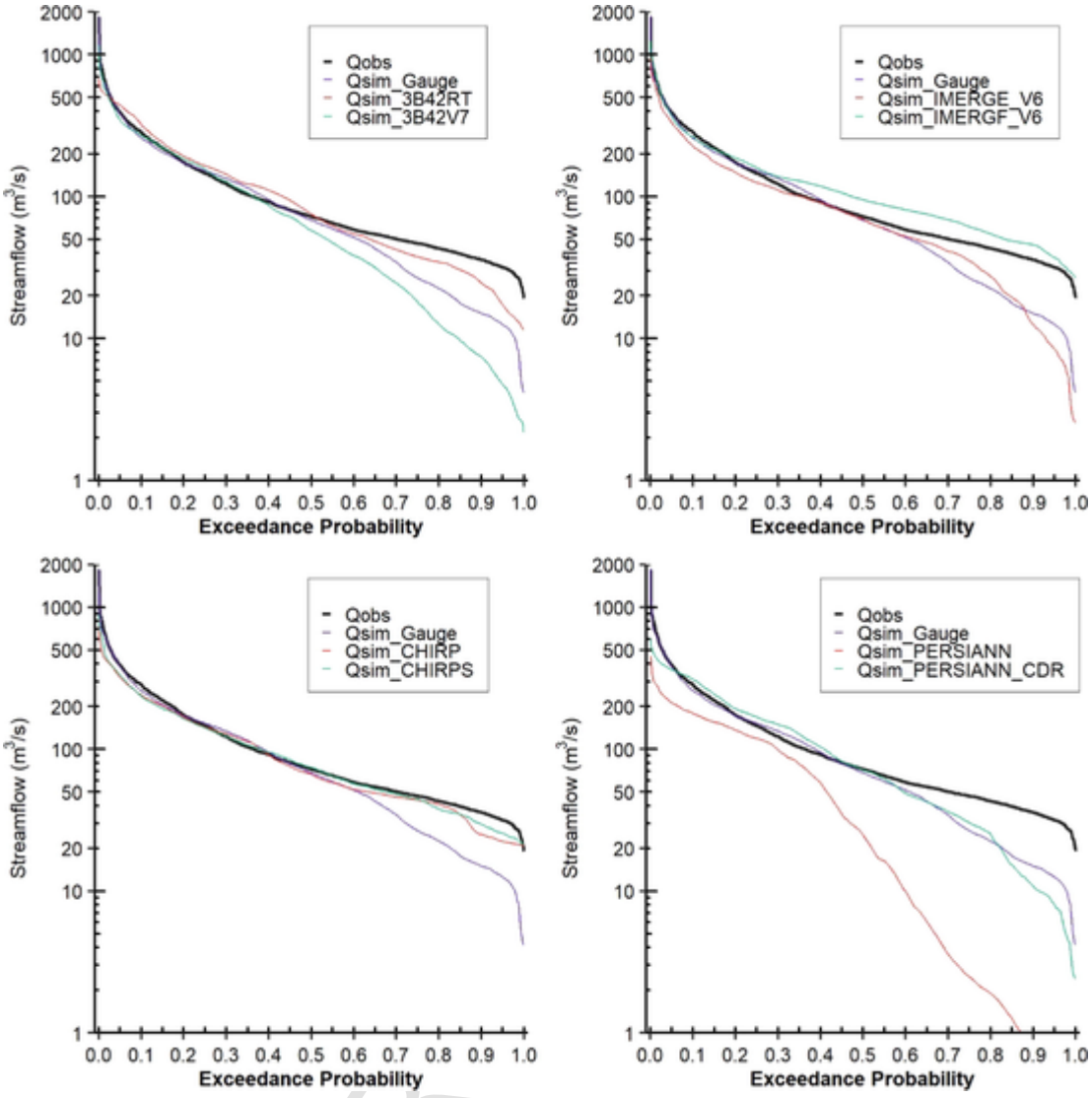

Fig. 11. Exceedance probability of the daily observed streamflow and simulated streamflow, driven by different precipitation inputs, at the XL basin, during the validation period (2010–2017). The logarithm was applied for the y-scale.

Declaration of Competing Interest

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

Acknowledgments

This research is funded by NASA project, under grant number NNX16AT86G, in collaboration with National Central for Water Resources Planning and Investigation (NAWAPI) within the framework of the project, grant number NE/S002847/1, funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED), and the project code VINIF.2019.DA17 funded by Vingroup Innovation Foundation (VINIF) annual research support program. We would like to express our sincere gratitude to the institutions for providing us with the required hydro-meteorological measurements that enabled this study. These include Vietnam Hydrology and Meteorological Administration, and National Central for Water Resources Planning and Investigation. Special acknowledgment also goes to NASA, University of California-Santa Barbara's Climate Hazards Group, and SERVIR-MEKONG for providing easy access to the SPE products and land cover data respectively. We would also like to thank Dr. R Srinivasan and Dr. Yihun Dele (Texas A&M Univ.) for their kind help and fruitful discussion regarding the application of SWAT, and Dr. Vu Manh Quyet (SFRI) for sharing soil map of Vietnam. We also acknowledge Mr. Truong Sinh Do (NAWAPI) for sharing the LS basin (Kycung River) dataset.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2020.124820.

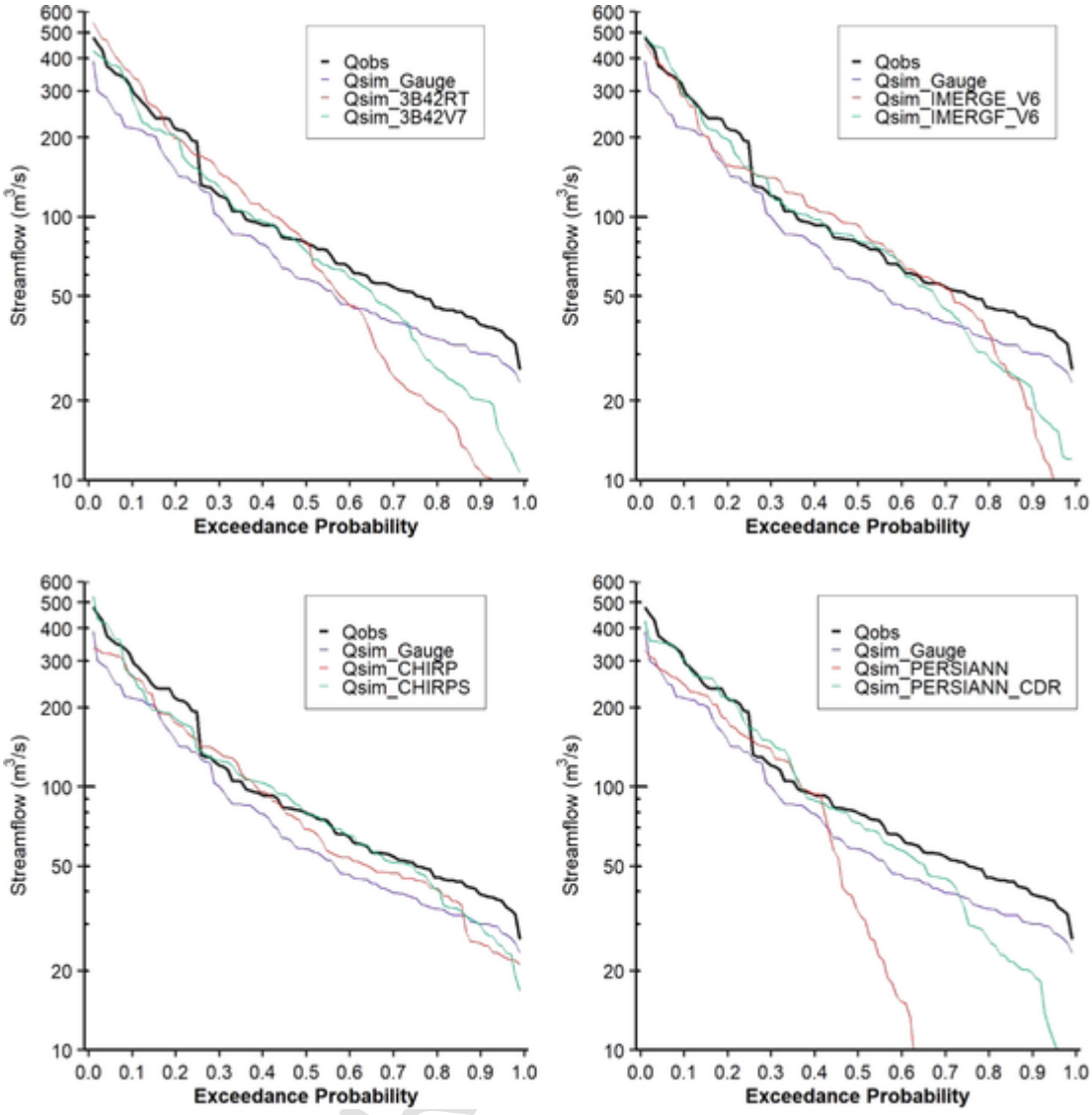

Fig. 12. Similar to Fig. 11 but for monthly dataset.

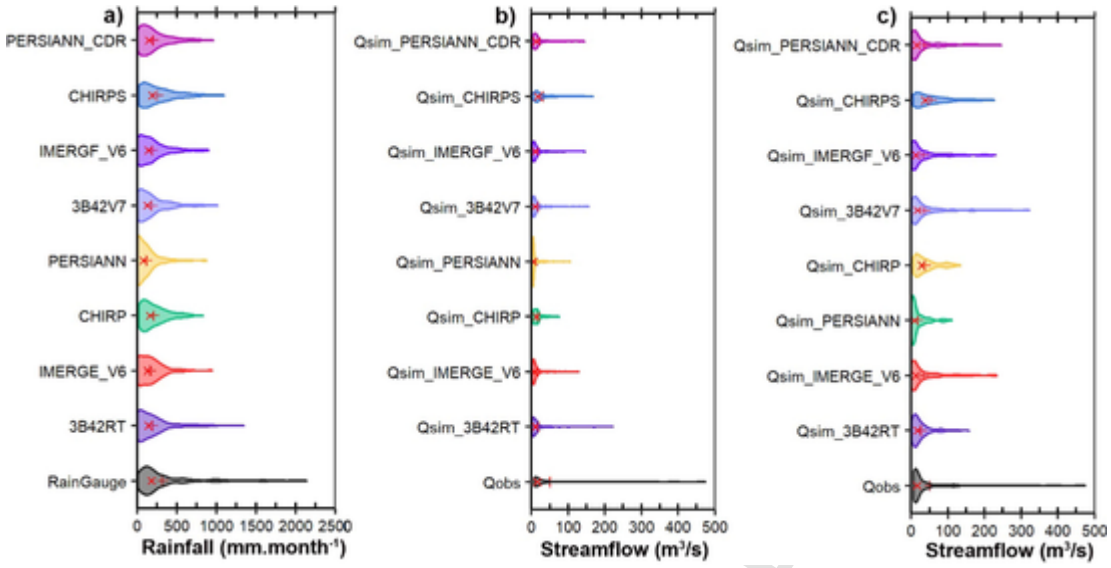

Fig. 13. Violin plots of a) monthly basin rainfall; b) streamflow simulation without re-calibration parameters (rain gauge parameters); c) streamflow simulation with re-calibration parameters using inputs from Satellite-derived Precipitation Estimation, at the AC basin. The cross sign indicates the median value; the plus sign indicates the mean value.

Table 4
The difference in median of precipitation and streamflow performance metrics between uncorrected and gauge-corrected SPE products. For all performance metrics, except for FAR, RMSE, and PBIAS, a positive value represents a better performance gauge-corrected version over its uncorrected version. The bold value indicates the gauge-corrected version worse than its uncorrected counterpart.

	Performances	3B42V7-3B42RT	IMERGF_V6-IMERGE_V6	CHIRPS-CHIRP	PERSIANN_CDR-PERSIANN
Precipitation Performance Metrics	POD	+0.121	+0.019	**−0.254**	+0.159
	FAR	−0.011	−0.012	−0.136	**+0.063**
	CSI	+0.031	+0.008	**−0.025**	+0.041
	CC	+0.114	+0.046	+0.032	+0.093
	RB	−0.1	+0.08	0.0	**+0.230**
	RMSE (mm. d^{-1})	−1.1	−0.4	−0.2	−0.6
Daily Streamflow Performance Metrics	NSE	+0.07	+0.12	+0.06	+0.24
	PBIAS (%)	−7.0	−7.6	−7.9	−18.7
Monthly Streamflow Performance Metrics	NSE	+0.20	+0.09	+0.17	+0.33
	PBIAS (%)	−4.7	−8.6	−6.4	−14.8

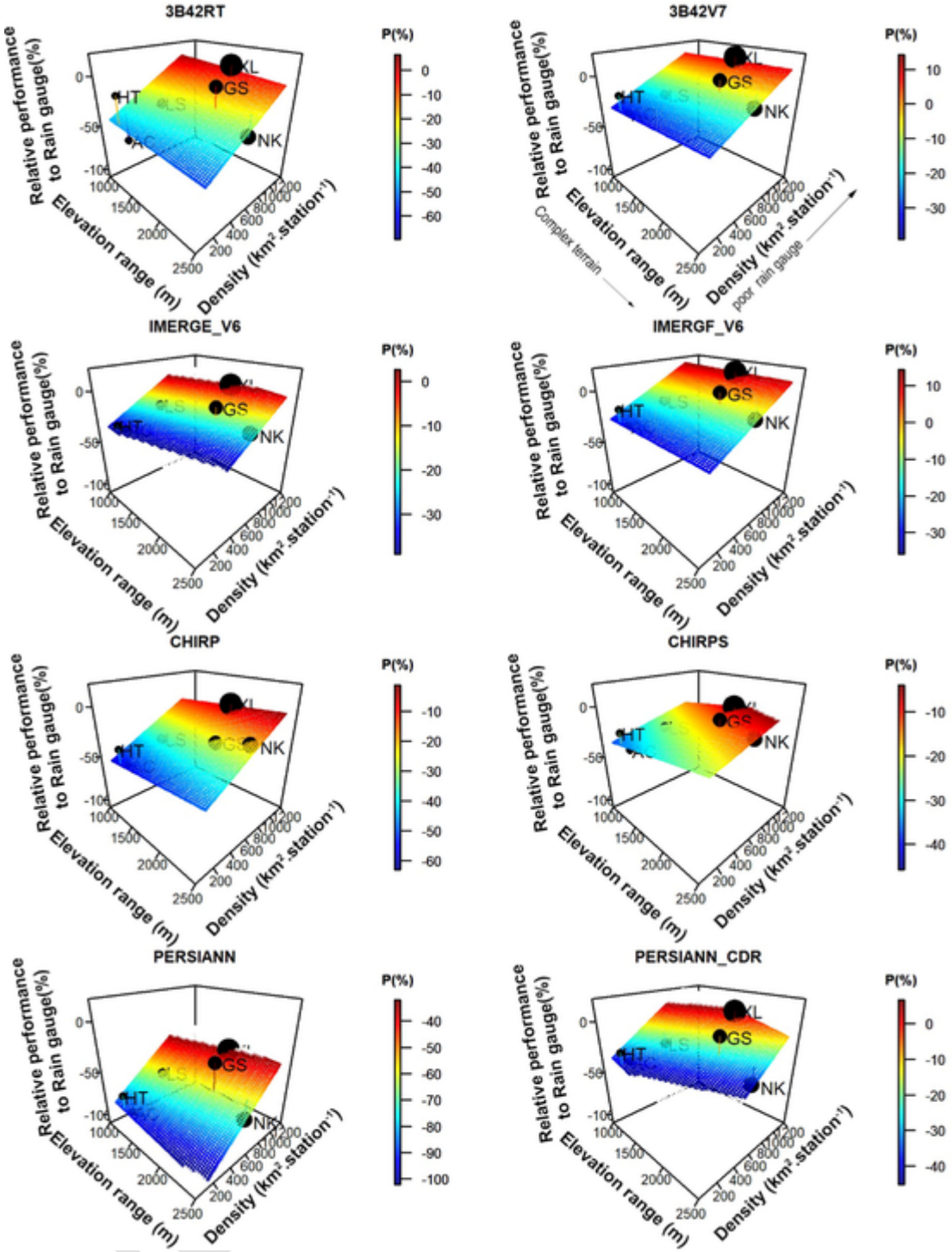

Fig. 14. Bivariate correlation analysis relative performance of SPE-driven simulations to rain gauge-driven simulation, elevation range, and rain gauge density. 3D surface denotes the performance's trend of SPE, compared to rain gauge (P%), as input for SWAT simulation. The black dot point indicates the relative size of the basin area.

References

Abbaspour, K C, et al., 2015. A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. J. Hydrol. 524, 733–752.

Abbaspour, K C, et al., 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. J. Hydrol. 333 (2–4), 413–430.

Alazzy, A A, et al., 2017. Evaluation of satellite precipitation products and their potential influence on hydrological modeling over the Ganzi river basin of the Tibetan Plateau. Adv. Meteorol..

Arnold, J G, Srinivasan, R, Muttiah, R S, Williams, J R, 1998. Large area hydrologic modeling and assessment part I: model development 1. JAWRA J. Am. Water Resour. Assoc. 34 (1), 73–89.

Ashouri, H, et al., 2015. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull. Am. Meteorol. Soc. 96 (1), 69–83.

Bai, L, Shi, C, Li, L, Yang, Y, Wu, J, 2018. Accuracy of CHIRPS satellite-rainfall products over mainland China. Remote Sensing 10 (3), 362.

Beck, H.E., et al., 2018. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. Hydrol. Earth Syst. Sci. Discuss.

Beck, H E, et al., 2017. MSWEP: 3-hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21 (1), 589.

Blöschl, G, 2013. Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales. Cambridge University Press.
Dile, Y T, Daggupati, P, George, C, Srinivasan, R, Arnold, J, 2016. Introducing a new open source GIS user interface for the SWAT model. Environ. Modell. Software 85, 129–138.
Duan, Z, et al., 2018. Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia. J. Hydrol..
Emmanuel, I, Andrieu, H, Leblois, E, Flahaut, B, 2012. Temporal and spatial variability of rainfall at the urban hydrological scale. J. Hydrol. 430, 162–172.
Funk, C, et al., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066.
Gerlak, A K, Lautze, J, Giordano, M, 2011. Water resources data and information exchange in transboundary water treaties. Int. Environ. Agreements: Politics, Law Econ. 2 (11).
Beck, H.E., et al., 2018. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. Hydrol. Earth Syst. Sci. Discuss.
Hargreaves, G H, Samani, Z A, 1982. Estimating potential evapotranspiration. J. Irrigation Drainage Division 108 (3), 225–230.
Hou, A Y, et al., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc. 95 (5), 701–722.
Huffman, G J, Bolvin, D T, 2013. TRMM and Other Data Precipitation Data Set Documentation. NASA, Greenbelt, USA, p. 28.
Huffman, G.J., et al., 2014. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm theoretical basis document, version, 4: 30.
Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2018. Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code, 612(2018): 47.
Huffman, G J, et al., 2007. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8 (1), 38–55.
Joyce, R J, Janowiak, J E, Arkin, P A, Xie, P, 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5 (3), 487–503.
Kidd, C, 2001. Satellite rainfall climatology: a review. Int. J. Climatol. 21 (9), 1041–1066.
Le, A, Pricope, N, 2017. Increasing the Accuracy of Runoff and Streamflow Simulation in the Nzoia Basin, Western Kenya, through the Incorporation of Satellite-Derived CHIRPS Data. Water 9 (2), 114.
Le, H, Sutton, J, Bui, D, Bolten, J, Lakshmi, V, 2018. Comparison and Bias Correction of TMPA Precipitation Products over the Lower Part of Red-Thai Binh River Basin of Vietnam. Remote Sensing 10 (10), 1582.
Li, D, Christakos, G, Ding, X, Wu, J, 2018. Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of Tiaoxi catchment (Taihu lake basin, China). J. Hydrol. 556, 1139–1152.
Li, Y, et al., 2019. Evaluation of three satellite-based precipitation products over the lower mekong river basin using rain gauge observations and hydrological modeling. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens..
Liu, X, Yang, T, Hsu, K, Liu, C, Sorooshian, S, 2017. Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. Hydrol. Earth Syst. Sci. 21 (1), 169.
Lobligeois, F, Andréassian, V, Perrin, C, Tabary, P, Loumagne, C, 2014. When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. Hydrol. Earth Syst. Sci. 18 (2), 575–594.
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. Geogr. Sci. 29 (1), 13–25.
Michaelides, S, et al., 2009. Precipitation: measurement, remote sensing, climatology and modeling. Atmos. Res. 94 (4), 512–533.
Mohammed, I N, Bolten, J D, Srinivasan, R, Lakshmi, V, 2018. improved hydrological decision support system for the Lower Mekong River Basin using satellite-based earth observations. Remote Sensing 10 (6).
Mondal, A, Lakshmi, V, Hashemi, H, 2018. Intercomparison of trend analysis of multisatellite monthly precipitation products and gauge measurements for river basins of India. J. Hydrol. 565, 779–790.
Moriasi, D N, Gitau, M W, Pai, N, Daggupati, P, 2015. Hydrologic and water quality models: performance measures and evaluation criteria. Trans. ASABE 58 (6), 1763–1785.
Nash, J E, Sutcliffe, J V, 1970. River flow forecasting through conceptual models part I—A discussion of principles. J. Hydrol. 10 (3), 282–290.
National Institute for Soils and Fertilizers, 2002. The basic information of main soil units of Vietnam. The Gioi Publishers, Hanoi, Vietnam.
NCHMF, 2000. Rainfall classification of Vietnam, Hanoi.
Neitsch, S L, Arnold, J G, Kiniry, J R, Williams, J R, 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009. Water Resources Institute.
Nguyen-Le, D, Matsumoto, J, Ngo-Duc, T, 2015. Onset of the rainy seasons in the eastern Indochina Peninsula. J. Clim. 28 (14), 5645–5666.
Nguyen-Xuan, T, et al., 2016. The Vietnam Gridded Precipitation (VnGP) Dataset: Construction and Validation. SOLA 12, 291–296.
Nguyen, D N, Nguyen, T H, 2004. Climate and Climate Resources in Vietnam (in Vietnamese). Agricultural Publishing House, Hanoi, Vietnam.
Nguyen, P, et al., 2019. The CHRS Data Portal, an easily accessible public repository for PERSIANN global satellite precipitation data. Sci. Data 6, 180296.
Nguyen, T H, Masih, I, Mohamed, Y A, van der Zaag, P, 2018. Validating Rainfall-Runoff Modelling Using Satellite-Based and Reanalysis Precipitation Products in the Sre Pok Catchment, the Mekong River Basin. Geosciences 8 (5), 164.
Ott, R L, Longnecker, M T, 2015. An Introduction to Statistical Methods and Data ANALYSIS. Nelson Education.
Phan, V-T, Ngo-Duc, T, 2009. Seasonal and interannual variations of surface climate elements over Vietnam. Clim. Res. 40 (1), 49–60.
Plengsaeng, B, Wehn, U, van der Zaag, P, 2014. Data-sharing bottlenecks in transboundary integrated water resources management: a case study of the Mekong River Commission's procedures for data sharing in the Thai context. Water Int. 39 (7), 933–951.
Rana, S, McGregor, J, Renwick, J, 2015. Precipitation seasonality over the Indian subcontinent: an evaluation of gauge, reanalyses, and satellite retrievals. J. Hydrometeorol. 16 (2), 631–651.
Ren, P, Li, J, Feng, P, Guo, Y, Ma, Q, 2018. Evaluation of multiple satellite precipitation products and their use in hydrological modelling over the Luanhe River Basin. China. Water 10 (6), 677.
Saha, S, et al., 2010. The NCEP climate forecast system reanalysis. Bull. Am. Meteorol. Soc. 91 (8), 1015–1058.
Sangati, M, Borga, M, 2009. Influence of rainfall spatial resolution on flash flood modelling. Nat. Hazards Earth Syst. Sci. 9 (2), 575–584.
Saxton, K E, Rawls, W J, 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci. Soc. Am. J. 70 (5), 1569–1578.
Sorooshian, S, et al., 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. Bull. Am. Meteorol. Soc. 81 (9), 2035–2046.
Thorndahl, S, et al., 2017. Weather radar rainfall data in urban hydrology. Hydrol. Earth Syst. Sci. 21 (3), 1359–1380.
Trinh-Tuan, L, Matsumoto, J, Ngo-Duc, T, Nodzu, M I, Inoue, T, 2019. Evaluation of satellite precipitation products over Central Vietnam. Prog. Earth Planet. Sci. 6 (1), 54.
Trinh-Tuan, L. et al., 2019b. Application of Quantile Mapping Bias Correction for Mid-future Precipitation Projections over Vietnam. SOLA.
Tuo, Y, Duan, Z, Disse, M, Chiogna, G, 2016. Evaluation of precipitation input for SWAT modeling in Alpine catchment: a case study in the Adige river basin (Italy). Sci. Total Environ. 573, 66–82.
USDA Soil Conservation Service, 1972. National Engineering Handbook Section 4 Hydrology, Chapters 4–10.
Ushio, T, et al., 2009. A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. J. Meteorol. Soc. Japan Ser. II 87, 137–151.
Vu, M, Raghavan, S V, Liong, S Y, 2012. SWAT use of gridded observations for simulating runoff–a Vietnam river basin study. Hydrol. Earth Syst. Sci. 16 (8), 2801–2811.
Vu, T T, Dao, N K, Do, Q L, 2017. Using gridded rainfall products in simulating streamflow in a tropical catchment–a case study of the Srepok River Catchment, Vietnam. J. Hydrol. Hydromech. 65 (1), 18–25.
Wang, W, Lu, H, 2016. Evaluation and Comparison of Newest GPM and TRMM Products over Mekong River Basin at daily scale. IEEE 613–616.
Wang, W, et al., 2016. Modelling hydrologic processes in the Mekong River Basin using a distributed model driven by satellite precipitation and rain gauge observations. PLoS ONE 11 (3), e0152229.
Williams, J R, 1969. Flood routing with variable travel time or variable storage coefficients. Trans. ASAE 12 (1), 100–0103.
Wu, Z, et al., 2018. Hydrologic evaluation of multi-source satellite precipitation products for the Upper Huaihe River Basin, China. Remote Sens. 10 (6), 840.
Zhang, Z, et al., 2019. Hydrologic Evaluation of TRMM and GPM IMERG Satellite-Based Precipitation in a Humid Basin of China. Remote Sens. 11 (4), 431.
Zhao, F, Zhang, L, Chiew, F H, Vaze, J, Cheng, L, 2013. The effect of spatial rainfall variability on water balance modelling for south-eastern Australian catchments. J. Hydrol. 493, 16–29.