Validating rainfall-runoff modelling using satellite-based and reanalysis precipitation products in the sre pok catchment, the mekong river basin

Nguyen, Thu Ha; Masih, Ilyas; Mohamed, Yasir A.; van der Zaag, Pieter

DOI
10.3390/geosciences8050164

Publication date
2018

Document Version
Final published version

Published in
Geosciences (Switzerland)

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
Validating Rainfall-Runoff Modelling Using Satellite-Based and Reanalysis Precipitation Products in the Sre Pok Catchment, the Mekong River Basin

Thu Ha Nguyen 1, Ilyas Masih 2,*, Yasir A. Mohamed 2,3 and Pieter van der Zaag 2,4

1 Southern Institute for Water Planning, Ho Chi Minh City 700000, Vietnam; haanhcmcat@gmail.com
2 IHE Delft Institute for Water Education, 2611 AX Delft, The Netherlands; y.mohamed@hrc-sudan.sd (Y.A.M.); p.vanderzaag@un-ihe.org (P.v.d.Z.)
3 Hydraulic Research Center, P.O. Box 318, Wad Medani, Sudan
4 Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2628 CD Delft, The Netherlands
* Correspondence: i.masih@un-ihe.org; Tel.: +31-152-152-340

Received: 15 February 2018; Accepted: 20 April 2018; Published: 4 May 2018

Abstract: The limitation in quantity and quality of climate data, in particular rainfall measurements, directly affects the reliability of hydrological models, which negatively impacts on the quality of water resources planning and management. This is obviously true in poorly-gauged or in transboundary river catchments with limited practice of data sharing among the riparian states. Satellite-based and reanalysis rainfall products are increasingly recognised as a promising information source to supplement ground observations. This research aims to study the feasibility of using satellite-based and reanalysis rainfall data for rainfall-runoff modelling in the Sre Pok catchment—a poorly gauged catchment located in Vietnam and Cambodia. First, three selected rainfall products (Tropical Rainfall Measuring Mission-TRMM, CPC MORPHing Technique-CMORPH, and European Reanalysis-ERA-Interim) were compared against ground rainfall data of 13 rain-gauges, located inside and in the vicinity of the study catchment. The validation results show that the studied global datasets have a reasonably good correlation with ground observation at monthly time scale (NSE = 0.3 to 0.8, $R^2 = 0.4$ to 0.8), but relatively poor correlation at daily scale (NSE = $-0.4$ to 0.3, $R^2 = 0.1$ to 0.4).

Next, a semi-distributed rainfall-runoff model based on HBV-Light was developed to validate these datasets in the Krong Ana sub-catchment of the Sre Pok. The results illustrate satisfactory to very good model performances at monthly time scale for the satellite-based and reanalysis rainfall inputs (NSE = 0.58 to 0.77, $R^2 = 0.59$ to 0.78 for the calibration period from 2004 to 2007, while NSE = 0.59 to 0.80, $R^2 = 0.64$ to 0.84 for the validation period from 2008 to 2010). However, only TRMM and ERA-Interim give acceptable results at daily scale with NSE = 0.52 to 0.67, $R^2 = 0.53$ to 0.67 for the calibration period, and NSE = 0.41 to 0.61, $R^2 = 0.45$ to 0.65 for the validation period. The study concludes that the tested satellite-based and reanalysis rainfall products, especially TRMM and ERA-Interim, could mimic reasonably well the monthly rainfall pattern in the study region. It also demonstrates a great opportunity to use satellite-based and reanalysis rainfall for runoff modelling at monthly time scale in the Sre Pok catchment and other poorly-gauged (transboundary) catchments in the region, for example in the Mekong and Red River Basins.

Keywords: rainfall-runoff; satellite-based rainfall; reanalysis rainfall; validation; Sre Pok catchment; Mekong basin; HBV
1. Introduction

Rainfall-runoff models offer a great opportunity to understand water resources system at the catchment scale [1,2], which is very important for water resources planning and management (IAHS 2013). However, developing a reliable rainfall-runoff model still remains a challenge due to the inadequacy and low accuracy of the input data, in particular precipitation estimates. Accurate precipitation measurement is still difficult due to its complex variability across time and space [1,3]. Traditionally, precipitation is monitored through ground-based measurement networks. The precipitation value at a station is considered to represent the average for a certain area around the station [4]. This results in inadequate representation of spatial variability of precipitation [1], which directly affects reliability of the rainfall-runoff models. This is true, especially for catchments with sparse or no rain gauge coverage [4–8].

In the last few decades, Satellite-Based and Reanalysis Precipitation (hereafter referred as SBRP) products has become a promising option to supplement Ground Measured Precipitation (hereafter referred as GMP) data [9,10] and facilitate researches on hydrology and modelling [11]. There has been a recognition of the advantages of satellite precipitation estimates [5] because of the high spatial and temporal resolution and the large areal coverage [12]. Thus, satellite-based and reanalysis precipitation estimates open the doors for improved hydrological assessment and water resources planning and management in poorly-gauged and un-gauged basins. This is in particular important for many regions in the world, where the density of precipitation gauges is normally low [6,10], or where data sharing is limited.

The application of different remotely sensed rainfall data in hydrological modelling has been studied for many years all over the world. However, there is no general conclusion on the accuracy and the reliability of streamflow simulation using these data sources. For example, Ref. [5] found that TMPA-3B42V6 results in simulated flows that were comparable to the ground observation at both 6-hour and monthly scales in Illinois river basin (USA). On the other hand, according to Ref. [13] this product cannot satisfactorily reproduce flows at daily scale in the Gilgel Abay watershed (the northwestern highland of Ethiopia). Similarly, Ref. [14] found that TRMM 3B42-V6 is inappropriate for daily flow simulation in Xinjiang catchment, Poyang lake basin, China. However, this rainfall product is successful in simulating the monthly flow in the same catchment. In other studies, Ref. [15] and Ref. [12], both demonstrated that flow simulation using satellite-based rainfall has a much better performance at monthly compared to daily time resolution in South America and the northeast of Tibet Plateau. In the upper Yellow river basin, daily TRMM-3B42V6 precipitation data could not simulate daily flow but good results were obtained when aggregated to monthly scale [12]. Ref. [16] suggested not using satellite-based rainfall estimates for low-flow simulation while almost all the tested satellite rainfall products result in good high-flow simulations in the 4 sub-basins of the Volta and Baro-Akobo basins. Ref. [17] also found that the satellite rainfall products will produce better results for the longer high-flow duration events.

There are some common lessons to be learned from the reviewed studies. Some authors agreed that model performance is improved when calibrated with respective satellite rainfall products compared to the calibration using observed rainfall (e.g., [13,16–19]). More promisingly, many scholars found a potential to use satellite-based rainfall sources for hydrological simulation (e.g., [5,13,15–20]. However, the quality of satellite-based rainfall products varies depending on the type of product, resolution (temporal and spatial), the season, climate, and topography of study areas [21–23]. For example, Ref. [23] found variable results with three precipitation products in case of seven basins of different characteristics in the southeastern United States. Therefore, a careful evaluation is required on case by case basis. This research aims to study the feasibility of using different SBRP data sources for a rainfall-runoff model to estimate runoff in the Sre Pok catchment (30,940 km²) of the Mekong River Basin, with focus on the Krong Ana sub-catchment (3080 Km²). First, the selected rainfall products were checked for their accuracy and reliability against GMP over the Sre Pok catchment. Next, a hydrological model was used to assess the accuracy of flow simulation using SBRP for the Krong Ana sub-catchment.
The paper consists of four main sections. Section 2 provides the information about the Sre Pok catchment and the selected rainfall-runoff model used for the evaluation. Section 2 also describes the required data and methods of preparing and validating them for rainfall-runoff simulation. The results and discussion are presented in Section 3. Finally, Section 4 synthesizes the summary and conclusions.

2. Materials and Methods

2.1. The Study Area

The Sre Pok catchment (30,940 km²) is a part of the Lower Mekong river basin, which lies on Vietnamese and Cambodian territory (Figure 1). Because of many reasons, mainly the poor resources to establish and maintain rain gauge network, the catchment has a very low density and uneven distribution of observation gauges, about 2380 km²/station (1513 km²/station in Vietnam and 12,780 km²/station in Cambodia), see Figure 1a. While data from Vietnam is available for a fairly long period and of relatively good quality, the data from Cambodia is poor and contains many gaps at all stations [24]. This situation will negatively affect reliability of hydrological modelling for the catchment, especially the sub-catchments without rain gauges.

![Figure 1. Study areas: (a) Sre Pok catchment; (b) Krong Ana sub-catchment; (c) location of the Sre Pok catchment in the Lower Mekong river basin; and (d) location of the Krong Ana sub-catchment in the Sre Pok catchment; TE: temperature and evaporation.](image)

The Sre Pok catchment has a wide range of topography with elevation varying between 56 m and 2409 m above mean sea level. Most of the catchment has low to medium elevation; only a small part in the southeast upper catchment is mountainous. The catchment is heavily influenced by tropical monsoon with two distinct seasons, rainy season from May to November and the dry season from December to April. Annual rainfall over the catchment ranges from 1400 to 2800 mm/year, which contributes about 32,000 Mm³/year to the total annual flow of the Mekong River [25]. However, the
flow has a very high variability over the year. Up to 87% of annual flow is generated from the rainy season. The average temperature of the catchment is about 22 °C, slightly cooler in Vietnam (21.2 °C) and slightly hotter in Cambodia (23.2 °C) [26].

2.2. Data Used

The ground-based daily rainfall data were collected for 13 stations from: (i) Post-Doctoral Research Programme on Adaptation to Climate Change, phase 2 (PRoACC-II, IHE Delft Institution for Water Education, Delft, The Netherlands), (ii) the Regional Flood Mitigation and Management Program (RFMMP, Mekong River Commission), and (iii) the Water Resources Section (WRS, Department of Agriculture and Rural Development, Dak Lak Province), Vietnam, as given in Table 1. The rainfall data from PRoACC-II and WRS were originally obtained from Vietnam’s Hydro-Meteorological Data Center and the Central Highland Region of Hydro-Meteorological Center (both are under Vietnam Ministry of Natural Resources and Environment). The data of RFMMP were provided by the National Mekong River Committee of Cambodia originated from the Department of Hydrology and River Works (Ministry of Water Resources and Meteorology of Cambodia, Phnom Penh, Cambodia). The first eight stations, PS01 to PS08, are located inside the Sre Pok catchment while the other five stations (PS09 to PS13) are located nearby to the catchment. Only one station (Stung Treng) is in Cambodian territory, see Figure 1a and Table 1. All the data were thoroughly checked for their quality using screening, time series graphs, and double mass analysis technique. The data were found to be complete and accurate in the common period of 2001 to 2010, except Stung Treng station (from 2004 to 2010). This research has selected three SBRP products: TRMM-3B42V7, CMORPH and ERA-Interim. Table 2 indicates selected studies where these data sets were used for hydrological modelling.

Table 1. Ground-based rainfall stations used in the research. In bold: stations used for the rainfall-runoff simulation of the Sre Pok sub-catchment.

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01</td>
<td>Krong Buk</td>
<td>108.4</td>
<td>12.8</td>
<td>456.7</td>
<td>PRoACC-II</td>
</tr>
<tr>
<td>P02</td>
<td>Duc Xuyen</td>
<td>108.0</td>
<td>12.3</td>
<td>444.6</td>
<td>PRoACC-II</td>
</tr>
<tr>
<td>P03</td>
<td>Giang Son</td>
<td>108.2</td>
<td>12.5</td>
<td>428.0</td>
<td>PRoACC-II</td>
</tr>
<tr>
<td>P04</td>
<td>Cau 14</td>
<td>107.9</td>
<td>12.6</td>
<td>310.1</td>
<td>WRS</td>
</tr>
<tr>
<td>P05</td>
<td>Ban Don</td>
<td>107.8</td>
<td>12.91</td>
<td>186.0</td>
<td>PRoACC-II</td>
</tr>
<tr>
<td>P06</td>
<td>Buon Ho</td>
<td>108.3</td>
<td>12.9</td>
<td>707.0</td>
<td>WRS</td>
</tr>
<tr>
<td>P07</td>
<td>BMT</td>
<td>108.1</td>
<td>12.7</td>
<td>472.1</td>
<td>WRS</td>
</tr>
<tr>
<td>P08</td>
<td>Dak Mil</td>
<td>107.6</td>
<td>12.4</td>
<td>759.6</td>
<td>WRS</td>
</tr>
<tr>
<td>P09</td>
<td>M’Drak</td>
<td>108.7</td>
<td>12.7</td>
<td>424.0</td>
<td>WRS</td>
</tr>
<tr>
<td>P10</td>
<td>Dak Nong</td>
<td>107.7</td>
<td>12.0</td>
<td>623.9</td>
<td>WRS</td>
</tr>
<tr>
<td>P11</td>
<td>Pleiku</td>
<td>108.0</td>
<td>14.0</td>
<td>780.3</td>
<td>WRS</td>
</tr>
<tr>
<td>P12</td>
<td>Da Lat</td>
<td>108.4</td>
<td>12.0</td>
<td>1511.0</td>
<td>WRS</td>
</tr>
<tr>
<td>P13</td>
<td>Stung Treng</td>
<td>106.0</td>
<td>13.5</td>
<td>52.6</td>
<td>RFMMP</td>
</tr>
</tbody>
</table>

These products were selected on the basis of their acceptable performance in previous studies, and because these data sets were publicly available. The TRMM (Tropical Rainfall Measuring Mission) is a joint satellite mission between NASA (The National Aeronautics and Space Administration of the USA, Washington, USA) and JAXA (the Japan Aerospace Exploration Agency, Tokyo, Japan). It was launched in November 1997 mainly for measurement of tropical and sub-tropical precipitation. The TRMM has three types of sensors including Precipitation Radar, Microwave Imager, and Visible Infrared Scanner. A wide range of TRMM products has been generated and can be freely accessed. TRMM-3B42V7 is the multiple-adjusted data product, available from 1 January 1998 for public assessment [3]. The product is available at relatively good spatial and temporal scales (0.25° × 0.25°, 3 h) and covers from 50° N to 50° S [27].

CMORPH (the Climate Prediction Center for morphing method) is the technique to generate rainfall data by combining motional vectors derived from geostationary satellite infrared data with precipitation obtained by passive microwave satellites. The rainfall product has a high temporal and
spatial resolution (30 min and 8 km). It covers the globe from 60° N to 60° S. The data is available at the resolution of 3 h, 0.25° × 0.25° on the website of Climate Predict Centre (NOAA, Silver Spring, MD, USA) from 3 December 2002 [28].

Table 2. Previous studies on hydrological modelling using satellite-based rainfall data.

<table>
<thead>
<tr>
<th>Products</th>
<th>Simulated Scale</th>
<th>Model Type</th>
<th>Catchment/Basin</th>
<th>Drainage Area (km²)</th>
<th>Location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMPA-3B42V6</td>
<td>Daily Monthly</td>
<td>Semi-distributed</td>
<td>7 sub-basin of La Plata basin</td>
<td>62,236–1,100,000</td>
<td>South America</td>
<td>Su et al. (2008)</td>
</tr>
<tr>
<td>CMORPH, TMPA-RT, TMPA-V6, PERSIANN, and PERSIANN-adjustment</td>
<td>6-h our Monthly</td>
<td>Lumped</td>
<td>Illinois River basin</td>
<td>1489</td>
<td>America</td>
<td>Behrangi et al. (2011)</td>
</tr>
<tr>
<td>CMORPH, TMPA-3B42RT, TMPA-3B42, and PERSIANN</td>
<td>Daily</td>
<td>Distributed</td>
<td>Gilgel Abay watershed</td>
<td>1656</td>
<td>Ethiopia</td>
<td>Bitew and Gebremichael (2011)</td>
</tr>
<tr>
<td>CMORPH, RFE2.0, TRMM-3B42, PERSIANN, and ERA-Interim</td>
<td>Daily</td>
<td>Distributed</td>
<td>4 sub-basins of Volta and Baro-Akobo basins</td>
<td>50,000–130,000</td>
<td>Africa</td>
<td>Thiemig et al. (2013)</td>
</tr>
<tr>
<td>TMPA 3B42V6, TMPA 3B42RT, and CMORPH</td>
<td>Daily Monthly</td>
<td>Distributed</td>
<td>Mishui basin</td>
<td>9972</td>
<td>South China</td>
<td>Jiang et al. (2012)</td>
</tr>
<tr>
<td>NOAA RFE2.0</td>
<td>Daily</td>
<td>Semi-distributed</td>
<td>4 sub-basins of Nile and Mekong rivers</td>
<td>2646–22,150</td>
<td>Eritrean, Ethiopia and Lao PDR</td>
<td>Artan et al. (2007)</td>
</tr>
</tbody>
</table>

The ECMWF (European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom), founded in 1975 is a research institute and weather service. ERA-Interim has been developed since 2001 and is the latest reanalysis atmospheric system in the world. It is the product of the numerical model based mainly on assimilated information of humidity and temperature [29]. The data is available from 1 January 1979 to 30 June 2014 with spatial and temporal resolutions of 0.25° × 0.25° and 3 h. However, the analyzed rainfall data is available only at 6 h time resolution.

The three datasets were downloaded, for the period of 2003–2010, in the format of netCDF through websites as mentioned in Table 3. The data were then transformed to the common time series data sets using scripts developed in Python language and aggregated into daily data. During this step, the difference of times between satellite and ground-based data was considered. The ECMWF, TRMM, and CMORPH use UTC time (Coordinated Universal Time) [28,30,31], whereas gauged rainfall is measured using local time. Due to national regulation of Vietnam, rainfall is recorded at two main times per day, 7:00 and 19:00 (GMT+7). The rainfall of today is the accumulated rainfall from 19:00 yesterday to 19:00 today. Daily satellite-based rainfall for today was accumulated from 12:00 of the previous day to 12:00 today. However, TRMM uses UTC centered time which is ±90 min offset from the nominal times [27]. For instance, the rainfall at 12:00 UTC was accumulated from 10:30 UTC to 13:30 UTC. Therefore, there was still a difference of 3 h in daily TRMM estimation compared to observation.

Table 3. Satellite-based and reanalysis products used in this research.

<table>
<thead>
<tr>
<th>No.</th>
<th>Type of Satellite Based Rainfall Product</th>
<th>Availability</th>
<th>Resolution</th>
<th>Downloading Period</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ECMWF (ERA-Interim)</td>
<td>Globally (1979-present)</td>
<td>0.25° × 0.25°</td>
<td>6 h</td>
<td>January 2003–December 2010</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>60N-60S (2003-present)</td>
<td>0.25° × 0.25°</td>
<td>3 h</td>
<td>January 2003–December 2010</td>
</tr>
<tr>
<td></td>
<td>TRMM (3B4-V7)</td>
<td>50N-50S (1998-present)</td>
<td>0.25° × 0.25°</td>
<td>3 h</td>
<td>January 2003–December 2010</td>
</tr>
</tbody>
</table>
2.3. Method of Data Comparison

Generally, there are three main methods to compare satellite-based (gridded data) and ground-based rainfall data, namely point-grid, areal, and grid-to-point comparison. The first is the direct comparison between gauged rainfall and rainfall of the respective grids in which gauges are located [32,33]. The second requires the interpolation of gauged data to be grid data; then the gauged-gridded data will be compared with grid data from remote sensing or re-analysis [12,13,32]. The third is the reversed method of the second, in that grid data are interpolated to the gauge cites by using bilinear interpolation of four adjacent-grid-cell values [34–36]. The second method is seen as a more meaningful comparison than the first as it compares exactly the same grids values while the first may involve large errors [13]. The second and the third demonstrate similar conclusions [34]. However, according to Ref. [35], the sparse-gauged network will significantly affect the accuracy of gauged-gridded data. Therefore, this study chose the third method using bilinear interpolation for the validation of satellite-based and reanalysis rainfall data.

After interpolation, all grid rainfall products were then compared to ground-measured rainfall at different temporal scales (daily and monthly) for the 13 stations during a common time period, from 2003 to 2010, except Stung Treng station, was from 2004 to 2010.

2.4. Rainfall-Runoff Model: HBV-Light

The HBV-Light model was chosen for the runoff simulation. The model has been developed by the Swedish Meteorological and Hydrological Institute (SMHI). HBV has been widely used for hydrological research [37–41]. The HBV model has been used for flood forecasting in Sweden since early 1970 [42]. The model has a simple and flexible structure and does not demand excessive datasets [40], which could be the reason for its wide application. In the Mekong region, the HBV model was applied in a study of ecosystem in Tonle Sap Lake [43], and developed for 13 Tonle Sap sub-catchments in a technical support project [44]. However, its application in the Mekong Region has not been common yet.

HBV is a conceptual rainfall-runoff model that simulates daily discharge using climatic inputs (daily rainfall, daily mean temperature, and daily potential evaporation). As many other hydrological models, the main principle of HBV is the water balance of the hydrological cycle. The model has four different routines including snow, soil moisture, response, and routing routines [37]. The model can be used as lumped or semi-distributed. For this research a Semi-distributed HBV model was developed, with different elevation bands and specific vegetation types representing the study catchment.

2.4.1. Model Setup

The rainfall-runoff model was intended to be developed for the entire Sre Pok catchment to verify the possibility of using SBRP for runoff simulation. However, the water discharge near the outlet has been modified because of the operation of hydropower plants in the upper mainstream. Moreover, with the limitation of data, it is difficult to set up a comprehensive model for the whole catchment. Therefore, the model was setup for a selected sub-catchment of Krong Ana where there is no influence of the hydropower plants, see Figure 1b,d. Its drainage area is about 3038 km$^2$, and elevation varies from about 399 m to 2408 m above mean sea level. The agricultural land covers approximately 50% of the total natural area while forest accounts for 15%; remaining area includes bare land and other land use. The observed annual rainfall in the sub-catchment varies between 1300 mm/year to 2420 mm/year. The average daily temperature is about 23.6 °C. Although the model allowed for detailed spatial representation of the Krong Ana catchment, the resolution of rainfall dataset (0.25° × 0.25°) is much coarser, e.g., the catchment has been covered by 5 to 6 pixels. This doesn’t allow detailed analysis of model sensitivity to spatial scale, but good enough to compare model results for similar resolutions of rainfall datasets.
Ground-based rainfall data from the six ground-based rainfall stations were used to develop the model (see Figure 1b and Table 1). Temperature data were collected from Water Resources Section, Department of Agriculture and Rural Development in Dak Lak province. The data sets were then checked for quality and data gaps were filled using regression analysis before preparing and using as inputs for the modelling purpose. Daily potential evaporation was calculated from maximum, minimum, and mean temperature using Hargreaves method [45]. The observed mean daily rainfall, temperature, and potential evaporation for the sub-catchment were estimated using Thiessen polygon method. The SBRP for the sub-catchment were calculated based on areal average method. The model was setup with 10 elevation bands and 3 vegetation zones.

2.4.2. Model Simulations

Considering the availability of both SBRP and GMP data, the period from January 2003 to December 2010 was selected for runoff simulation. The year 2003 was used for warming-up the model; 2004–2007 for calibration, while 2008–2010 was used for validation. The calibration and validation were conducted separately for GMP and for each of the SBRP (TRMM, CMORPH and ERA-Interim). In fact, first we have attempted to use the same calibration parameters from the model run with MGP input, and apply it for the other datasets. The results were poor. Then, we followed automatic calibration for each model separately. This is in agreement with analysis given in the literature (e.g., [13,16,18,19]).

All results were compared with each other to assess the accuracy of runoff simulation using different types of SBRP inputs. During calibration, the Genetic Algorithm optimization method in automatic calibration was applied to find the most suitable parameters set for each simulation. The Genetic Algorithm was chosen on the basis that this method was deemed efficient and required fewer simulations compared to other optimization methods, such as Monte-Carlo method, which is also available in the HBV-Light Software. For example, trial tests indicated acceptable results with 6000 simulations using Genetic Algorithm when compared with the results of over 100,000 simulations by Monte-Carlo. The parameter ranges for optimization were based on Ref. [40].

The Nash-Sutcliffe Efficiency (NSE) was used as an objective function. The best parameter set is the one that produces maximum value of the objective function. The performance of the best parameter set was also evaluated using additional evaluation criteria, discussed below.

Daily flow at the outlet of the sub-catchment (Giang Son station) was used as a reference flow for calibration and validation.

2.4.3. Evaluation Indicators

To evaluate the accuracy of SBRP data compared to the ground observation and the performance of the model, both visual inspection and statistical criteria such as NSE, coefficient of determination ($R^2$), and Bias (%) were applied. Besides, NSE with logarithmic values (LogNSE) was also employed for evaluating the model performance.

\[
\text{NSE} = 1 - \frac{\sum (Q_{obs}(t) - Q_{sim}(t))^2}{\sum (Q_{obs}(t) - \bar{Q}_{obs})^2} \quad (1)
\]

\[
\text{LogNSE} = 1 - \frac{\sum (lnQ_{obs}(t) - lnQ_{sim}(t))^2}{\sum (lnQ_{obs}(t) - ln\bar{Q}_{obs})^2} \quad (2)
\]

\[
R^2 = \frac{(\sum (Q_{obs}(t) - \bar{Q}_{obs})(Q_{sim}(t) - \bar{Q}_{sim}))^2}{\sum (Q_{obs}(t) - \bar{Q}_{obs})^2 \sum (Q_{sim}(t) - \bar{Q}_{sim})^2} \quad (3)
\]

\[
\text{Bias} = \frac{\sum (Q_{sim}(t) - Q_{obs}(t))}{\sum Q_{obs}(t)} \times 100 \quad (4)
\]
where, $Q_{obs}(t)$ and $Q_{sim}(t)$ are the observed and simulated discharge at time $t$, $Q_{obs}$ and $Q_{sim}$ are the average observed and simulated discharge.

Ref. [46] reviewed 128 cases of flow simulation and suggested general ranges of statistical indicators for model performance. According to the authors, a simulation can be ranked as satisfactory when NSE > 0.50 and bias < ±25%, as good performance when 0.65 < NSE ≤ 0.75 and ±10% ≤ bias < ±15%, and as very good when NSE > 0.75 and bias ≤ ±10%. These ranges were applied for the performance evaluation of this study. Also, in terms of R2 and LogNSE, the research used the same performance ratings as NSE.

3. Results and Discussion

3.1. Validation of SBRP against GMP

First the performance of rainfall products for the daily time scale is presented. The studied SBRP products generally overestimate the number of rainy days per year at most of the stations (Figure 2a). TRMM and CMORPH have similar trends and perform much better than ERA-Interim. The number of rainy days by ERA-Interim is almost double to the observed values at ground stations. This result is similar to the finding of Ref. [33] in the Zambezi and Volta basins in Africa. According to the authors, this might be due to the inability of sensors to differentiate between drizzle and light rain. However, this might also reflect the ability of sensors to capture areal rainfall in the surrounding area of a ground rainfall station when there is no rain at that station [47]. Similarly, the maximum daily rainfall is also overestimated by the examined SBRP products (Figure 2b). For average rainfall over the Krong Ana Sub-catchment, the trend is quite different. Only Era-Interim significantly overestimates the number of rainy days and the value of average maximum daily rainfall. TRMM and CMORPH demonstrate a performance that slightly underestimates these factors. TRMM agrees well with GMP.

![Figure 2](image_url)  
**Figure 2.** (a) Comparison of average number of rainy days per year of different Satellite-Based and Reanalysis Precipitation (SBRP) and Ground Measured Precipitation (GMP) data sets; (b) The ability of SBRP in capturing the average maximum daily rainfall.

At daily time step, the estimated values of $R^2$ and NSE (Table 4) illustrate that none of the SBRP agrees well with the GMP. The correlations between them are very poor. The values of $R^2$ vary from 0.11 to 0.39, and most of the lowest values are for ERA-Interim (from 0.11 to 0.26). The $R^2$ values of TRMM are slightly higher than that of CMORPH and ERA-Interim over all stations, except for Dak Mil (P08). In terms of NSE, CMORPH demonstrates best values among the examined SBRP products.
Table 4. Statistical indicators for the comparison between SBRP and GMP data sets.

<table>
<thead>
<tr>
<th>ID</th>
<th>Daily Indicators</th>
<th>Monthly Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( R^2 ) TRMM</td>
<td>( NSE ) TRMM</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>ERA-Interim</td>
</tr>
<tr>
<td>P01</td>
<td>0.39</td>
<td>0.14</td>
</tr>
<tr>
<td>P02</td>
<td>0.26</td>
<td>0.12</td>
</tr>
<tr>
<td>P03</td>
<td>0.30</td>
<td>0.17</td>
</tr>
<tr>
<td>P04</td>
<td>0.28</td>
<td>0.16</td>
</tr>
<tr>
<td>P05</td>
<td>0.31</td>
<td>0.16</td>
</tr>
<tr>
<td>P06</td>
<td>0.32</td>
<td>0.19</td>
</tr>
<tr>
<td>P07</td>
<td>0.28</td>
<td>0.17</td>
</tr>
<tr>
<td>P08</td>
<td>0.26</td>
<td>0.13</td>
</tr>
<tr>
<td>P09</td>
<td>0.36</td>
<td>0.20</td>
</tr>
<tr>
<td>P10</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>P11</td>
<td>0.34</td>
<td>0.24</td>
</tr>
<tr>
<td>P12</td>
<td>0.26</td>
<td>0.14</td>
</tr>
<tr>
<td>P13</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>K.An</td>
<td>0.49</td>
<td>0.37</td>
</tr>
</tbody>
</table>
However, all the tested products have generally very low NSE values with the maximum of 0.33. At sub-catchment scale, the values of $R^2$ and NSE for Krong Ana are also low although higher than those at all stations of the three SBRP products. As shown in Figure 3, for mean daily rainfall, TRMM overestimates rainfall at most stations of the sub-catchment. ERA-Interim tends to underestimate rainfall while CMORPH shows a mix pattern at all stations. However, both CMORPH and ERA-Interim underestimate mean rainfall for Krong Ana sub-catchment. However, there is also an underestimated trend evident from all products at M’Drak, Dak Nong and Pleiku station (P09, P10 and P11).

Figure 3. Validation of SBRP against GMP at daily scale: (a) Mean daily rainfall and (b) bias (%) at different stations.

The results of the comparison of mean monthly rainfall are shown in Figure 4. The figure illustrates that the seasonal pattern is well reflected by all rainfall sources. However, three SBRP products overestimate the rainfall in the dry months (November to April, see Figure 5), while showing a relatively better agreement during the wet months (May to September, see Figure 6). The abnormal values are generally larger for the dry season than for the wet season. ERA-Interim has a tendency to underestimate the rainfall at all stations whereas TRMM and CMORPH show variable patterns.

The result of statistical indicators at the monthly timescale is more promising and far better compared to the daily time scale as given in Table 5. About two thirds of $R^2$ are greater than 0.65, which reflect a good linear correlation between SBRP and GMP. TRMM has the best correlation at most of the stations with $R^2$ ranging from 0.58 to 0.81, and 11 over 13 stations having an $R^2$ greater than 0.65. ERA-Interim has slightly better $R^2$ values compared to CMORPH in most cases. $R^2$ values for ERA-Interim fluctuate between 0.52 and 0.74 while those of CMORPH range between 0.4 and 0.82. CMORPH has better NSE values than ERA-Interim. On the whole, TRMM demonstrates the best performance at most of the stations.
The results of the comparison of mean monthly rainfall are shown in Figure 4. The figure illustrates that the seasonal pattern is well reflected by all rainfall sources. However, three SBRP products overestimate the rainfall in the dry months (November to April, see Figure 5), while CMORPH has better NSE values than ERA-Interim. On the whole, TRMM demonstrates the best performance at most of the stations.

1. Comparison of average monthly rainfall patterns (2003–2010) from different data sources at different stations and Krong Ana Sub-catchment: (a) at Krong Buk station (P01), (b) at Giang Son station (P03), (c) at Buon Ho station (P06), (d) at BMT station (P07), (e) at M’Drak station (P09), (f) at Dat Lat station (P12), and (g) at Krong Ana Sub-catchment.
Figure 5. Monthly bias (%) of different SBRP data sets compared to GMP data in dry season at different stations and Krong Ana Sub-catchment: (a) at Krong Buk station (P01), (b) at Giang Son station (P03), (c) at Buon Ho station (P06), (d) at BMT station (P07), (e) at M’Drak station (P09), (f) at Dat Lat station (P12), and (g) at Krong Ana Sub-catchment.
These results are consistent with the validation of daily rainfall given in the previous section. Validation periods most of the time, simulations using TRMM tend to overestimate the water flow more than 20%, while those based on TRMM data decrease to less than 10%. In both calibration and simulations of daily flows achieve reasonable results during calibration period, with biases in the extent though they have relatively low skill to capture high peaks over both calibration and validation (2008–2010). Also, all simulations using SBRP data are able to imitate the runoff to some extent.

Although simulated flow based on GMP is lower than observed flow, the differences are insignificant, results of the simulations using different rainfall sources in comparison with the observed flow at the outlet of the Krong Ana sub-catchment. The graphs show the calibration results during 2004 to 2007 and validation results during 2008 to 2010. The performance indicators are given in Table 5.

The hydrographs of Figure 7 show that the simulated flow from ground rainfall generally has the best fit with observed flow although the simulation slightly underestimates the largest peaks during calibration (2004–2007) and overestimates them during validation periods. The generated flow from TRMM has the best agreement with observed flow among the three SBRP products. ERA-Interim and CMORPH are likely to simulate incorrectly the occurrence of monthly bias (%) at Buon Ho (P06)

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6}
\caption{Monthly bias (%) of different SBRP data sets compared to GMP data in rainy season at different stations and Krong Ana Sub-catchment: (a) at Krong Buk station (P01), (b) at Giang Son station (P03), (c) at Buon Ho station (P06), (d) at BMT station (P07), (e) at M'Drak station (P09), (f) at Dat Lat station (P12), and (g) at Krong Ana Sub-catchment.}
\end{figure}
3.2. Runoff Simulations

3.2.1. Runoff Simulation at Daily Timescale

We first present the results of rainfall-runoff simulation at daily timescale. Figure 7 presents the results of the simulations using different rainfall sources in comparison with the observed flow at the outlet of the Krong Ana sub-catchment. The graphs show the calibration results during 2004 to 2007 and validation results during 2008 to 2010. The performance indicators are given in Table 5.

The hydrographs of Figure 7 show that the simulated flow from ground rainfall generally has the best fit with observed flow although the simulation slightly underestimates the largest peaks during calibration (2004–2007) and overestimates them during validation (2008–2010). Also, all simulations using SBRP data are able to imitate the runoff to some extent though they have relatively low skill to capture high peaks over both calibration and validation periods. The generated flow from TRMM has the best agreement with observed flow among the three SBRP products. ERA-Interim and CMORPH are likely to simulate incorrectly the occurrence of highest peaks. However, none of the simulations (including simulation using gauged rainfall) was able to capture the highest peak in December 2005. This peak occurred in the middle of December, which was considered an abnormality in the Krong Ana sub-catchment as well as the Sre Pok catchment. The normal highest peaks usually occur in early November or October or even earlier in August.

Comparing the mean daily flows during calibration and validation period supports the above results on rainfall evaluation as depicted by Figure 8. The most comparable simulated runoff to observed values is the runoff derived by the GMP data, which is better than the three SBRP products. Although simulated flow based on GMP is lower than observed flow, the differences are insignificant, only −3.6% for calibration and −0.5% for validation. Deriving the model with the SBRP inputs, all simulations of daily flows achieve reasonable results during calibration period, with biases in the range of ±15%. However, for the validation, the bias from CMORPH and ERA-Interim increase to more than 20%, while those based on TRMM data decrease to less than 10%. In both calibration and validation periods most of the time, simulations using TRMM tend to overestimate the water flow (bias > 0) while simulations using CMORPH and ERA-Interim tend to underestimate the flow (bias < 0). These results are consistent with the validation of daily rainfall given in the previous section.

3.2.2. Runoff Aggregation at Monthly Timescale

In general, the monthly aggregated flows from daily simulations have better agreements with the observed flow, which is shown by the good fits between simulated and observed hydrographs in Figure 9. The aggregated monthly flow simulated using GMP data still shows best results with $R^2$ of 0.93 and 0.96, NSE of 0.92 and 0.95, logNSE of 0.92 and 0.92 for calibration and validation respectively. There is a significant improvement in the performance of all three SBRP products, compared to the daily timescale, which is exhibited by the statistical indicators presented in Table 5, TRMM performing best. Although the ability of TRMM to reproduce low flow is lower than that of ERA-Interim (logNSE of TRMM < logNSE of ERA-Interim), overall it achieved a very good performance while ERA-Interim ranked good. Accumulated monthly flow using CMORPH data was acceptable. TRMM sometimes overestimates the high peaks (in 2004, 2009 and 2010) while ERA-Interim and CMORPH have a main tendency to underestimate them. TRMM captures high peaks fairly well whereas ERA-Interim and CMORPH sometimes fail to capture them, for example in 2005, 2008 and 2009.
Figure 7. Hydrographs comparing simulated and observed flow at daily time scale based on rainfall input data from (a) GMP; and the SBRP; (b) Tropical Rainfall Measuring Mission (TRMM); (c) Climate Prediction Center for morphing method (CMORPH); and (d) ERA-Interim for calibration period (January 2004–December 2007) and validation period (January 2008–December 2010).
results are poor for all three products with NSE < 0.33 and R² < 0.39. Besides, all three SBRP data sets tend to overestimate rainfall in the dry months (December to April) and the last two months of rainy throughout the year as well as rainfall during the rainy months (May to September). However, they whereas ERA-Interim and CMORPH sometimes fail to capture them, for example in 2005, 2008 and Interim (logNSE of TRMM < logNSE of ERA-Interim), overall it achieved a very good performance performing best. Although the ability of TRMM to reproduce low flow is lower than that of ERA-

Table 5. Statistical criteria of different simulations at daily scale.

<table>
<thead>
<tr>
<th></th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GMP</td>
<td>TRMM</td>
</tr>
<tr>
<td>R²</td>
<td>0.89</td>
<td>0.67</td>
</tr>
<tr>
<td>NSE</td>
<td>0.89</td>
<td>0.67</td>
</tr>
<tr>
<td>LogNSE</td>
<td>0.89</td>
<td>0.62</td>
</tr>
<tr>
<td>Bias (%)</td>
<td>-3.61</td>
<td>10.93</td>
</tr>
<tr>
<td>Daily simulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly accumulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.93</td>
<td>0.78</td>
</tr>
<tr>
<td>NSE</td>
<td>0.92</td>
<td>0.77</td>
</tr>
<tr>
<td>LogNSE</td>
<td>0.92</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Figure 8. (a) Average simulated vs. average observation daily flow and (b) Bias between simulated flow from different rainfall sources and measured flow (blue: calibration period, red pattern: validation period).

Figure 9. Hydrograph comparing average simulated and observed flow at monthly scale for calibration period (January 2004–December 2007) and validation period (January 2008–December 2010).

The comparison between simulated flows from the SBRP and GMP shows that the GMP is the best input in the Krong Ana sub-catchment. The flow generated from the GMP has a better fit of hydrograph with observed flow and has much better values of the examined performance indicators compared to simulated flows from SBRP inputs. This tendency is the same for both daily flow and aggregated monthly flow. The very good result of simulation with GMP is understandable when this simulation used rainfall data from six stations over an area of 3038 km². The density of rainfall station in this case is about 506 km²/station and therefore within the requirement for hilly and mountainous area (250–575 km²/station [48]. However, many other sub-catchments in the Sre Pok
lack sufficient ground-based rainfall stations. Many located in Cambodian territory do not have any rainfall station. In this case, satellite-based and reanalysis rainfall estimates would bring significant benefits for rainfall-runoff assessment to these sub-catchments.

4. Summary and Conclusions

The main objective of this research was to study the feasibility of using different SBRP products for a reliable rainfall-runoff simulation in the Sre Pok catchment. Three SBRP products, namely TRMM, CMORPH, and ERA-Interim were selected for this purpose. The three products were first evaluated against GMP at thirteen stations using bilinear interpolation method to examine their accuracy as input for a rainfall-runoff model. The semi-distributed HBV model was then selected to test the runoff simulation using satellite-based and reanalysis rainfall for Krong Ana, an upstream sub-catchment of the Sre Pok catchment. To evaluate the performance of the satellite and reanalysis rainfall products, four main statistical indicators, NSE, logNSE, $R^2$, and bias (%), were considered.

The results show that all satellite and reanalysis products have a reasonably good correlation with ground observation at monthly timescale but relatively poor correlation at daily scale. Although CMORPH has a slightly better performance than TRMM and ERA-Interim at daily scale, the statistical results are poor for all three products with NSE < 0.33 and $R^2$ < 0.39. Besides, all three SBRP data sets considerably overestimate the number of rainy days per year. At monthly scale, most of the NSE and $R^2$ values are larger than 0.5, about 44% of NSE and 67% of $R^2$ values being greater than 0.65. In general, TRMM shows the best performance among the three datasets with NSE of 0.47 to 0.75 and $R^2$ of 0.58 to 0.81. All products demonstrate a high ability to correctly capture rainfall patterns throughout the year as well as rainfall during the rainy months (May to September). However, they tend to overestimate rainfall in the dry months (December to April) and the last two months of rainy season (October and November).

The modelling results demonstrate a good potential to apply SBRP data for runoff simulation in the Krong Ana sub-catchment. Generally, the simulated flows from the tested SBRP data sources give a reasonable resemblance to daily observed flow over the calibration period with $R^2$ > 0.53, NSE > 0.52, and bias < ±15%. The results are a bit lower for the validation period with only satisfactory performance for TRMM ($R^2 = 0.65$, NSE = 0.61, bias = 7.32%) and ERA-Interim ($R^2 = 0.52$, NSE = 0.49, bias = −20.75%) inputs. The simulation driven by TRMM has the best performance among the three rainfall products though it tends to overestimate the average daily water flow in the sub-catchment. The simulations using CMORPH and ERA-Interim are likely to underestimate the water flow. All simulations with the studied SBRP products show a relatively low skill to capture the magnitude of high peaks over both calibration and validation periods. The comparison between monthly flows obtained from the SBRP and GMP inputs shows a significant improvement, which is evident from estimated performance indicates and a better agreement in hydrographs for both calibration and validation periods. TRMM input yields the best model performance at monthly scale with $R^2$ of 0.78 and 0.84 and NSE of 0.77 and 0.80 for calibration and validation respectively. Although ERA-Interim performs much better than CMORPH during calibration, they have similar results for validation. Both of them result in satisfactory model performance at this scale ($R^2$ and NSE > 0.5).

Overall, this study demonstrates that SBRP data can be a valuable source of rainfall estimates at monthly scale, and as input for runoff simulation in the study sub-catchment of Krong Ana. The encouraging results from this research show a good opportunity to use SBRP for runoff modelling in the whole Sre Pok catchment or similar poorly gauged catchments of the Mekong Basin. The research only evaluated the possibility of using three satellite and reanalysis rainfall products for runoff simulation while in fact there are many other rainfall products and data from public domain sources, for example temperature, evaporation, wind speed, or solar energy. The use of these SBRP data for runoff simulation in poorly-gauged or un-gauged catchments is worth studying.

Author Contributions: T.H.N. completed data processing and analysis, model simulation, result analysis, and drafting of the manuscript. I.M. provided expertise in assessing indicators, model parameters and set up, and
completed the manuscript. Y.A.M. and P.v.d.Z. contributed to the research proposal, methods, result analysis and review the manuscript.

Acknowledgments: The research was supported by the Netherland Fellowship Program and IHE Delft (formerly UNESCO-IHIE) Institute for Water Education. We are thankful for the warm welcome, support, and data provision during field visit and data collection of Pham Tan Ha, the Department of Agriculture and Rural Development of Dak Lak Province, Vietnam and Nguyen Quoc Anh, the Regional Flood Mitigation and Management Programme—the Mekong River Commission, Phnom Penh, Cambodia. We also would like to extend our gratefulness to Schalk Jan van Andel (UNESCO-IHIE), Isnaeni Hartanto (UNESCO-IHIE) and Tim Hessels (Delft University of Technology) for their help and useful advice in processing of satellite-based and reanalysis rainfall data.

Conflicts of Interest: The authors declare no conflict of interest.

References


7. Li, M.; Shao, Q. An improved statistical approach to merge satellite rainfall estimates and raingauge data. *J. Hydrol.* 2010, 385, 51–64. [CrossRef]


23. Yilmaz, K.K.; Hogue, T.S.; Hsu, K.-L.; Sorooshian, S.; Gupta, H.V.; Wagener, T. Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeorol.* 2005, 6, 497–517. [CrossRef]


40. Masih, I.; Uhlenbrook, S.; Maskey, S.; Ahmad, M.D. Regionalization of a conceptual rainfall–runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran. *J. Hydrol.* **2010**, *391*, 188–201. [CrossRef]


