Available online at www.CivileJournal.org

Civil Engineering Journal

Vol. 5, No. 1, January, 2019

Evaluation of Multi-Precipitation Products for Multi-Time Scales and Spatial Distribution During 2007-2015

Nguyen Tien Thanh [a*]

[a] *Department of Hydro-meteorological Modeling and Forecasting, Thuyloi University, 175 Tay Son, Dong Da, Hanoi, Vietnam.*

Received 29 October 2018; Accepted 06 January 2019

Abstract

Recently, several precipitation products are released with the improved algorithm to strengthen the performance of precipitation construction and monitoring. These data play a key role in a wide range of hydrological models, water resources modeling and environmental researches. Especially in developing countries like Vietnam, it is challenging to gather data for long-term time series at scales of daily and sub-daily due to the very coarse density of observation station. In order to overcome the problem of data scarcity, this study aims to evaluate the performance of newest multiple precipitation products including Tropical Rainfall Measuring Mission (TRMM 3B42 V7), Climate Prediction Center (CPC) MORPHing Version 1.0 (CMORPH_V1.0), European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis systems (ERA-Interim), Climate Research Unit Time series Version 4.0.1 (CRU TS 4.0.1) and Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources version 2 (APHRODITE) in comparison with measured precipitation for multiple time scales (daily, monthly, seasonal and annual), taking the VuGia-ThuBon (VG-TB) as a pilot basin where climate regime is complex. Seven continuous and four dichotomous statistics are applied to evaluate the precipitation estimates qualitatively at multiple time scales. In addition, specifically, evaluation of spatial distribution of multiple time scales is implemented. The results show lower precipitation estimates in areas of high elevation and higher precipitation estimates over the areas of plain and coastal in comparison with measured precipitation for all considered precipitation data. More importantly, ERA-Interim well captures rain events of heavy rain (50.0-100 mm/day). CMORHPH_V1.0 better reproduces the rain events with little overestimation of light rain (0.6-6 mm/day) than the others. For zero rain events (0-0.6 mm/day), TRMM 3B42 V7 gives the best performance. Furthermore, the cumulative distribution function of APHRODITE well matches the distribution of measured precipitation. All precipitation products completely fail to capture the rain events of extremely heavy rain. More importantly, a formula is proposed to scale and adjust the merged satellite precipitation at a sub-daily scale.

Keywords: CMORPH; TRMM; ERA-Interim; Satellite Precipitation; Gridded Precipitation; VuGia-ThuBon.

1. Introduction

Precipitation is considered to be one of the key inputs for the fields of hydrological and environmental. In the field of meteorological, precipitation is an important term of the energy budget, but a challenge of high quality parameterization within weather and climate models (e.g., convection precipitation parameterization schemes). The reason for this can come from a variety of factors (e.g., the limitation of the observation system, geography, or uneven distribution of precipitation in space and time). A classification of precipitation data can be divided into gridded gauge data (e.g., Asian Precipitation -Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) [1], Climate Research Unit Time Series (CRUTS)), satellite data (e.g., Cics High-resolution Optimally Interpolated

* Corresponding author: thanhnt@tlu.edu.vn

http://dx.doi.org/10.28991/cej-2019-03091242

Microwave Precipitation From Satellites (CHOMPS) and merged satellite data (e.g., Tropical Rainfall Measurement Mission (TRMM), Global Satellite Mapping of Precipitation (GSMaP) [2]). Alternatively, precipitation data is categorized into the data of weather station (e.g., Global Historical Climatology Network Daily Temperatures (GHCND) [3], Global Surface Summary of the Day (GSOD)), reanalysis (e.g., Global Land Data Assimilation System (GLDAS)) [4] and gridded gauge data (e.g., APHRODITE, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), Climate Prediction Center (CPC) MORPHing (CMORPH) [5]). This is valuable data sources for hydrological and environmental modelling. However, each type of data is assimilated by different methods and set a given spatio-temporal resolution. For instance, APHRODITE is set a 0.25 x 0.25oresolution of grid spacing for daily over the monsoon Asia [1]. Furthermore, the accurate level of data varies in spatio-temporal scales from region to region. Therefore, a quality assessment of precipitation products on multi-time scales of spatio-temporal is especially important not only for end-users but for developers.

In practice, it is difficult to construct a high quality of gauge-based datasets because of the inhomogeneous nature of the source data. Plus, wind and objects (e.g., trees) have significantly effect on the amount of collected water (e.g., site on the wind trajectories). Examples of gauge-based datasets are the Climatic Research Unit Time series(CRU TS) the University of the East Anglia (UEA) and Global Precipitation Climate Centre (GPCC) [6]. The CRU TS data is reconstructed over land areas as much of a global scale as possible from over 4000 weather stations around the world [7]. Meanwhile, the GPCC data is merged by the Deutscher Wetterdienst, the National Meteorological Service of Germany [6]. Although observing systems and algorithms are updating over the years, the quality and potential uncertainty of each data are a big challenge for researchers [8-10]. For mountains of Vietnam, specially, the coarse density of observation station is very low. For this, several distribution (e.g., Gamma or Johnson) are introduced [11, 12]. These distributions, however, should be further investigated for different regions. Hence, an evaluation of precipitation data needs to be considered.

Parallel to the development of remote sensing technologies and computer systems, satellite products are very useful information sources for the earth science studies. Some of advantages are high spatio-temporal resolution over large areas, global coverage, and continuous severe weather measurement. As compared with gauge-based data, however, atmospheric, hydrological, and oceanographic parameters are not directly measured. The missing data may derive from the sensor failure. Furthermore, most satellites operated over a region only twice per day give a potential missing data. For this, a merged satellite data is often implemented based on the instruments of microwave and/or infrared. As any observing systems, it is very difficult to evaluate the accuracy of satellite retrievals. The reason for this is the effects of different factors (e.g., cloud, atmospheric absorption, topography, aerosols) [13, 14]. To date, even a new technology merges the precipitation from several satellite-based algorithms and rain gauges (i.e., the CPC Merged Analysis of Precipitation), it is important to evaluate the precipitation estimations. Fekete et al[15] compared and assessed the precipitation datasets of the CRU, Willmott-Matsuura (WM), GPCC, Global Precipitation Climatology Project (GPCC), Tropical Rainfall Measuring Mission (TRMM), and NCEP–Department of Energy (DOE) Atmospheric Model Intercomparison Project (AMIP-II) Reanalysis (NCEP-2). The potential uncertainty in these datasets was illustrated for arid and semiarid regions. Examples of studying in evaluation of precipitation datasets can be also found in the publications [10, 16-20]. Recently, evaluation of satellite-based rainfall products is done for different domains over the world (e.g., Tramblay et al. [21] for Marocco, Kenabatho et al. [22] for Botswana or Ayehu et al. [23] for Ethiopia). Most studies, however, are performed at the coarse resolution of temporal (e.g., monthly) for a large scale of spatial with one or very few precipitation products. Furthermore, climate regime in Vietnam is extremely complex as precipitation regime is strongly governed by the tropical monsoon and trade winds, geographical conditions, moisture source from the sea and encroachment of cold fronts from the Siberian high pressure.

To interpret the causes of flood events, the studies of precipitation features is considered to be a central input. Precipitation at a scale of sub-daily is importantly required. At this scale, however, the precipitation data is not broadly available. Vietnam as a special example, a series of precipitation data at coarser levels (e.g., daily) is mainly provided from the terrestrial observational system (e.g., meteorological station and rain gauge). Furthermore, over Vietnam, there are more than 2000 daily recording stations, but just over 100 hourly stations exist. There is an overall density of about one station per 1500 km2, 2500 km2 for the north and south of Vietnam, respectively [24].This is a great obstruction for the studies in meteorology, hydrology and environment. For this, the satellite-related datasets from the global and regional observing systems provide very valuable information on the weather processes. Each data, however, is set at a given resolution of spatio-temporal. It is especially noted that more recent precipitation estimates from a span of different sources have upgraded and improved for tropical areas (e.g., APHRODITE-2 products [1]). So, an evaluation of newest precipitation products is urgent to find the best product for a given region, especially in Vietnam where exists lots of ungauged or poorly gauged regions as well as under an abnormal changes in meteorological forcings under a global warming.

In this study, the purpose is to evaluate the newest multi-precipitation products (i.e., TRMM 3B42 v7, CMORPH_V.1.0, APHRODITE, CRU TS4.0.1 and ERA-Interim) at multi-time scales and spatial distribution, taking

VuGia-ThuBon basin located in central Vietnam as an example. The results are validated against observed precipitation time series during a 9-years period (2007-2015).

2. Materials and Methodology

2.1. Study Area

In this study, Vu Gia-Thu Bon (VG-TB) basin is selected. It is located in central Vietnam, elongating from 16o55’ through 14o55’ and from 107o15’ through 108o24’ and covers a total of area of approximately 12577 km2. The VG-TB basin is surrounded by two main provincial administrative territories Quang Nam and Da Nang. The basin is characterized by a steep topography and the altitude ranging from 0 m at the coast to 2567 m in elevation in the west (Figure 1)

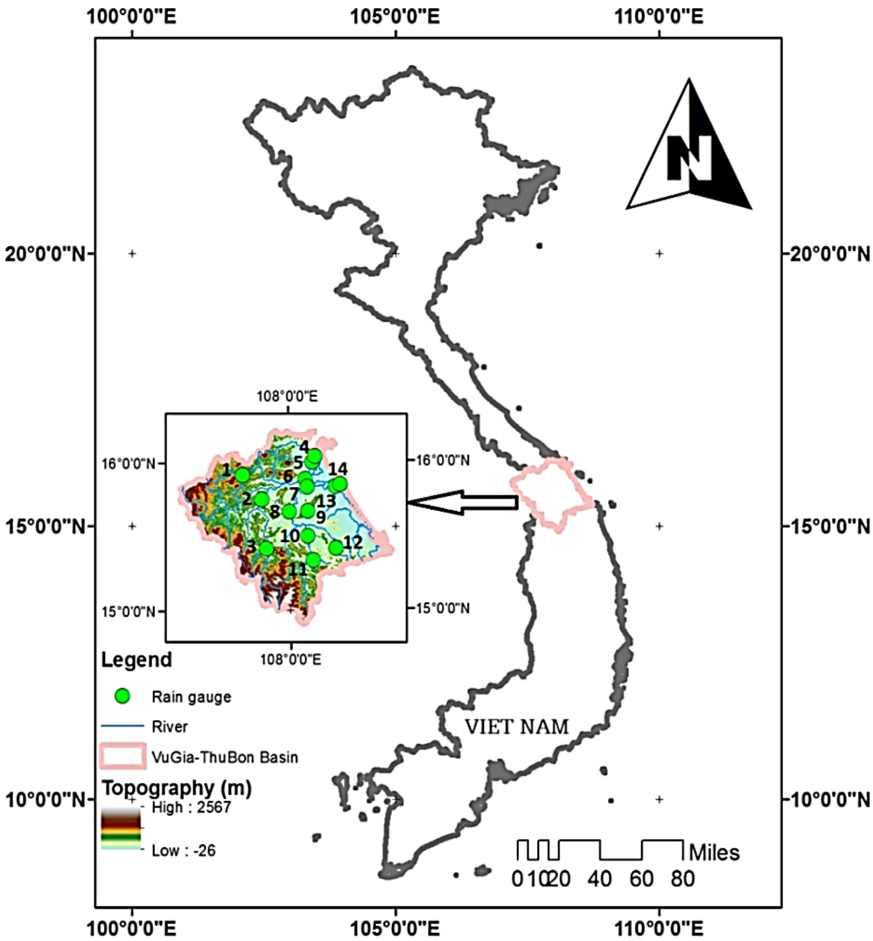

Figure 1. Network of hydro-meteorological stations at Vu Gia-Thu Bon basin in central Vietnam (1: Hien, 2: Thanh My, 3: Kham Duc, 4: Da Nang, 5: Cam Le, 6: Ai Nghia, 7: GiaoThuy, 8: Nong Son, 9: Que Son, 10: HiepDuc, 11: Tra My, 12: Tien Phuoc, 13: Cau Lau, 14: Hoi An)

2.2. Data

2.2.1. Terrestrial Observation Datasets

The daily precipitation records are obtained from the Vietnam HydroMeteorological Data Center, which belongs to the Ministry of National Resources and Environment of Vietnam (MONRE). The monthly, seasonal and annual precipitation records are aggregated from the daily series of data. There are two national rain gauge stations (i.e., Danang and Tramy). Other stations including Ainghia, Camle, Giaothuy, Caulau, Hien, Hiepduc, Hoian, Khamduc, Nongson, Queson, Thanhmy and Tienphuocare popular rain gauge stations which operate manually on the base of the volunteer of local people. Location of these stations is displayed on Figure 1. The data is available from 2007-2015.

2.2.2. Merged Satellite, Gridded Gauge and Reanalysis Datasets

TRMM 3B42 V7: The TRMM is a joint space mission between the NASA and Japan's National Space Development Agency. A variety of instruments are used namely the TRMM Microwave Imager (TRMM TMI), the TRMM Precipitation Radar (TRMM PR), Clouds and Earths Radiant Energy System, Visible Infrared Scanner and Lightning

Imaging Sensor. Mainly, the estimation of precipitation is based on the TRMM TMI and TRMM PR. These data is processed using an algorithm in order to obtain the TRMM Combined Instrument calibration data set (TRMM 2B31) for the TRMM Multi-satellite Precipitation Analysis (TMPA), the TMPA 3B43 monthly precipitation and the TMPA 3B42 daily and sub-daily (3-hours). In this study, a product of TRMM 3B42 V7 is evaluated [25]

CMORPH_V.1.0: Precipitation data is constructed at very high spatio-temporal resolution from low or biter satellite microwave observations exclusively, and whose features are transported via spatial propagation information that is obtained entirely from geostationary satellite IR data [5]. The algorithms of Ferraro [26], Ferraro et al. [27] and Kummerow et al. [28] are used to estimate precipitation from the passive microwaves aboard the DMSP 13, 14&15 (SSM/I), the NOAA-15, 16, 17 & 18 (AMSU-B), and AMSR-E and TMI aboard NASA's Aqua and TRMM spacecraft, respectively. Precipitation estimations are available on a grid of 27 km, even finer of 8 km in dealing with spatial resolution. In temporal scale, daily, 3-hours, and 30-minutes are available. In this study, the version of CMORPH_V1.0 is used to evaluate at a spatio-temporal resolution of 3-hourly and 0.25 degree.

APHRODITE: Since 2006, daily high-resolution grids of precipitation are produced for the entire Asia within APHRODITE's Water Resources project, which is conducted by the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of the Japan Meteorological Agency (MRI/JMA). The datasets cover the period 1951-2007. Data is based on a rain gauge observational network. The density of the rain gauges network is of higher density over the Himalayas, South and Southeast Asia and mountainous areas in the Middle East. The number of gauges varies between 5 000 and 12 000 and is therefore two and four times higher than those that are considered by WMO GTS [1].

CRU TS 4.0.1: Established in 1972, the Climatic Research Unit (CRU) at the University of the East Anglia (UEA) is reconstructed the measured data for climate data over land areas as much of a globe scale as possible. These datasets currently have a resolution of 0.5o x 0.5o latitude/longitude degrees. They are developed from the observed data from over 4 000 weather stations around the world [7]. Almost all these weather stations are exchanged via the National Meteorological Services (NMSs) under the WMO's sponsorship; the data from these stations are exchanged via the CLIMAT network, which is part of the WMO Global Telecommunications System (GTS). In the early 1980s, many of the original datasets are published in decadal data publications entitled 'World Weather Records, WWR'. They also publish data from the National Climatic Data Centre in Asheville, North Carolina, USA. Both the gridded datasets and the station data archived have evolved over the years. The latest versions of CRU datasets are available via the British Atmospheric Data Centre (BADC) consisting of CRU TS 3.22, CRU TS 3.23, CRU TS 3.24, CRU TS 4.0.0 and CRU TS 4.0.1. The CRU TS is monthly gridded data based on daily values; the ASCII and NetCDF files are available for various parameters such as mean temperature and radiation.

ERA-Interim: This is a third generation reanalysis of the global atmosphere, initiated in 2006 to reproduce the previous reanalysis the European Centre for Medium-Range Weather Forecasts (ECMWF with the main objective to improve the ERA-40. The ERA-40 (1957-2002) is a second-generation reanalysis. The term "ERA" refers to the computerized weather data from the European Centre for Medium-Range Weather Forecasts. The "reanalysis" stands for a method that creates a comprehensive picture of the state of the Earth system on a four-dimensional grid. Generally, this method is a combination of weather prediction model and observations to produce gridded datasets of many meteorological fields (e.g., temperature or precipitation) at a high resolution of temporal. The products created with this method are known as reanalysis data. The ERA-Interim atmospheric model and reanalysis system are configured for (1) a spatial resolution of 60 levels in the vertical with the top level at the 0.1 hPa, (2) T255 spherical-harmonic representation for the basic dynamical fields and (3) a reduced Gaussian grid with a horizontal resolution of approximately 0.7 degrees latitude/longitude for surface and other grid-point fields [29].The data assimilation system used to produce ERA-Interim is based on a 2006 release of the IFS (Cy31r2). The system includes a 4-dimensional variation analysis (4D-Var) with a 12-hour analysis window. The spatial resolution of the data set is approximately 80 km (T255 spectral) on 60 vertical levels from the surface up to 0.1 hPa [30].

Table 1. Sources of data

	TRMM 3B42 V7	**CMORPH_V1.0**	**APHRODITE**	**CRU TS 4.0.1**	**ERA-Interim**
Geographic coverage	50°S - 50°N	Global (60N - 60S)	Monsoon Asia (60E - 150E, 15S - 55N)	Global land	Global
Temporal resolution	3-Hourly	3-Hourly	daily	Monthly	3-hourly
Horizontal resolution	0.25° x 0.25°	0.25° x 0.25°	0.25° x 0.25°	0.5° x 0.5°	0.125°x 0.125°
Period	1998-2018	2002-2018	1951-2018	1901-2016	1979-2018
File type	HDF	NetCDF	NetCDF	NetCDF	GRID

1.1. Methodology

In order to enable consistent point-to-pixel and pixel-to-point comparisons, all precipitation of merged satellite, gridded gauge and reanalysis datasets, hereafter refers as considered precipitation estimates, are upscaled and downscaled to a unified grid of 0.125◦. The up scaling and downscaling procedure applied in this study consisted in transferring values from the high-resolution raster cells to each one of the 0.125◦ grid cells, by using technology of bilinear interpolation.

2.3.1. Visual Verification

One of the oldest and best verification methods is the good old fashioned visual, or "eyeball”, method. In this study, we examine the difference between the considered precipitation estimates and measured precipitation. Visual verification is an instantaneous judgment to distinguish the error between the considered precipitation estimates and measured precipitation. We often use exploratory graph techniques such as spatial distribution, time series plot, histogram and Cumulative Distribution Function (CDF) referring to the use of visual verification.

2.3.2. Continuous Statistics

To evaluate the performance of considered precipitation estimates, eight continuous statistics are used in this study. Bias is defined as the average difference between measured precipitation and considered precipitation estimates. A negative bias indicates underestimation of rainfall while a positive bias indicates overestimation. A perfect value of estimate would result in a bias of 0. The multiplicative bias (Mbias) is the ratio of considered datasets to rain gauge value with a perfect score of 1. Values less than 1 indicate underestimations, while values greater than 1 refer to overestimations. The mean absolute error (MAE) is used to represent the average magnitude of the error with a perfect score of 0. The root mean square error (RMSE), which gives a greater weight to the larger errors relative to MAE, is used to measure the average error magnitude. The correlation coefficient (r) is used to measure how close the points of a scatter plot are to a straight line. It refers the agreement between considered precipitation estimates and measured precipitation. The Nash–Sutcliffe efficiency with logarithmic value (ln(Nash)) and Nash–Sutcliffe efficiency (Nash) are given. The Nash–Sutcliffe efficiency with logarithmic values ln(Nash) is selected because it can be added to expect a better quantification of the performance in different conditions (maximum and minimum values). The formulas and a brief description of these statistics are given as follows:

Table 2. Continuous statistics

Index	Formula	Range
RMSE	$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)^2}$	$0 \le \mathrm{RMSE} < \infty$ Perfect score: 0
Bias	$\mathrm{Bias} = \frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)$	$-\infty < \mathrm{Bias} < \infty$ Perfect score: 0
MAE	$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\lvert S_i - O_i \rvert$	$-\infty < \mathrm{Bias} < \infty$ Perfect score: 0
Correlation coefficient	$r = \frac{\sum_1^N (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (S_i - \bar{S})^2}\sqrt{\sum_{i=1}^N (O_i - \bar{O})^2}}$	$-1 \le r \le 1$ Perfect score: 1
ln(Nash)	$\ln(\mathrm{Nash}) = 1 - \frac{\sum_1^n (\ln O_i - \ln S_i)^2}{\sum_1^n (\ln O_i - \ln \bar{S})^2}$	$-\infty \le \ln(\mathrm{Nash}) < 1$ Perfect score: 1
Nash	$\mathrm{Nash} = 1 - \frac{\sum_1^n (O_i - S_i)^2}{\sum_1^n (O_i - \bar{S})^2}$	$-\infty \le \mathrm{Nash} < 1$ Perfect score: 1
Mbias	$\mathrm{Mbias} = \frac{\sum_{i=1}^N S_i}{\sum_{i=1}^N O_i}$	$-\infty < \mathrm{Bias} < \infty$ Perfect score: 1

Where S_i is the value of considered precipitation estimates for the i^{th} daily event, O_i is the value of measured precipitation for the i^{th} daily event, N is the number of precipitation events, $\overline{S_i}$ is the average valueof considered precipitation product for N daily events over each grid box, and $\overline{S_i}$ is the average value of measured precipitation for N daily events over each grid box.

2.3.3. Dichotomous Statistics

The Threat score (critical success index-denoted CSI), Probability of detection (hit rate-denoted POD), Probability of false detection (false alarm rate-denoted FAR) and Probability of false alarm detection (false alarm rate-denoted POFD) are used to assess the rain-detection capabilities of considered precipitation estimates. CSI measures the fraction of measured precipitation that is correctly diagnosed by the considered precipitation products. POD indicates the ratio of the correct identification of precipitation by considered precipitation estimates to the number of precipitation occurrences observed by measured precipitation. FAR presents the proportion of cases in which the considered precipitation records rainfall when the terrestrial observations do not. POFD is the ratio between the false alarms to the non-gauge observed events.

Table 3. Dichotomous statistics

Index	Formula	Range
CSI	$\mathrm{CSI} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}}$	$0 \leq$ CSI≤ 1 Perfect score: 1
POD	$\mathrm{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}}$	$0 \leq$ POD≤ 1 Perfect score: 1
FAR	$\mathrm{FAR} = \frac{\text{hits}}{\text{hits} + \text{misses}}$	$0 \leq$ FAR≤ 1 Perfect score: 1
POFD	$\mathrm{POFD} = \frac{\text{false alarms}}{\text{correct negatives} + \text{flase alarms}}$	0≤POFD≤1 Perfect score: 0

Where hits represent the number of times that observed rain is correctly detected, misses indicate the number of times that observed rain is not detected, correct negatives indicate the number of times that rain is not detected but not observed and false alarms are the number of times that rain is detected but not observed. It is noted that a threshold of 0.6 mm/day is used to distinguish between rain and no rain.

We further analyzed the skills of considered precipitation estimates at estimating various precipitation event types by comparing their distributions of daily rainfall rates to those recorded by the rain gauges. In presenting our results, we adopted the following precipitation classification criteria (mm/day): zero rain, 0–0.6; light rain, 0.6–6.0; moderate rain, 6.0-16.0; heavy rain, 16.0–50.0, very heavy rain, 50.0–100, and extremely heavy rain, above 100. We computed the probability of occurrence of each precipitation type from the entire time series for each merge satellite-, reanalysis- and gridded gauge–measured precipitation observation pair. For each region and precipitation class, the statistics are summarized in a boxplot to represent all data pairs, and the probability distributions are compared between the rain gauge, TMPA version 6, and TMPA version 7 datasets.

3. Results and Discussions

3.1. Visual Verification

As the first step of visual verification, a comparison between measured precipitation and considered precipitation estimates is implemented at a various span of tempo-spatial distribution. Temporal scales of annual, seasonal, monthly and daily precipitation are aggregated from precipitation of 3-hours (TRMM 3B42 V7, CMORPH_V1.0, ERA-Interim), daily (APHRODITE, measured precipitation) and monthly (CRU TS 4.0.1). In this study, four seasons of winter (December, January and February), spring (March, April and May), autumn (September, October and November) and summer (June, July and August) are investigated. Only typical two seasons (winter and summer) are analysed and presented as displayed on Figure 3.

Figure 2 shows the spatial distribution of precipitation datasets for annual (a) and daily (b) temporal scales over VG-TB basin. Compared to measured precipitation, precipitation products produced by merged satellite of CMORPH_V1.0 and TRMM 3B42 V7 tend to a very underestimation of precipitation. Opposed to this, reanalysis data of ERA-Interim produces precipitation much higher than measured precipitation over VB-TB basin. An explanation for this is that a single or ensemble of weather patterns (i.e., tropical storms, polar fronts and intertropical convergence zone) are mainly causes of heavy rainfall events for short time durations. Specially, over the central Vietnam involved the VG-TB basin is strongly dominated by the tropical monsoon of a peninsula in the Southeast of the European-Asian continent and trade winds. In addition, the effect of topographical shapes of Bach Ma and Truong Son mountain series located in the north and west of basin significantly contribute to uneven distribution of precipitation and complex of rainfall regime. All these could lead to highly failed estimations. Besides, these may relate to the TRMM 3B42 v7 and CMORPH_V1.0 retrieval algorithms as well as the physical aspects of the reanalysis system and the method of data assimilation of ERA-Interim. So, these factors should be carefully considered in the reproduction of heat and moisture as well as formation of precipitation. It is specially noted that the gridded data of CRU TS 4.0.1 and APHRODITE represent better performances at scales of both annual and daily with an underestimation. In general, the precipitation datasets of merge satellite, gridded and reanalysis do not capture the spatial pattern of precipitation as compared to the measured precipitation. The best precipitation estimations refer to APHRODITE, followed by CRU TS 4.0.1.

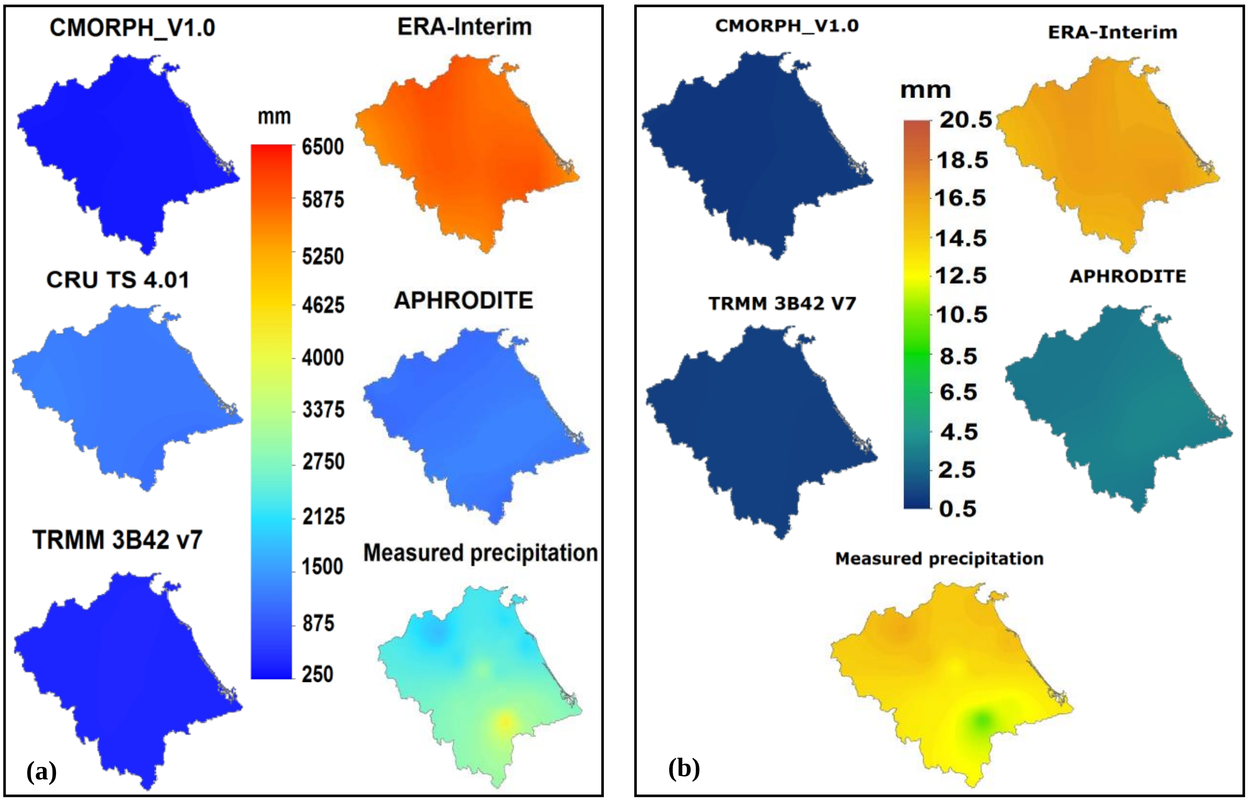

Figure 2. Spatial distribution of precipitation datasets for annual (a) and daily (b) temporal scales during 2007 - 2015 over VG-TB basin

A wide span of precipitation is found out over VG-TB basin for CMORPH_V1.0 (329-397 mm/year), CRU TS 4.0.1 (1120-1304 mm/year), TRMM 3B42 V7 (422-448 mm/year), ERA-Interim (5470-6148 mm/year), APHRODITE (1074-1331 mm/year) and measured precipitation (1845-4170 mm/year). For the daily scale, precipitation widely ranges for CMORPH_V1.0 (0.9-1.1 mm/day), TRMM 3B42 V7 (1.2-1.3 mm/day), ERA-Interim (14.9-16.9 mm/day), APHRODITE (2.9-3.8 mm/day) and measured precipitation (5.1-11.4 mm/day). For the seasonal scale, it can be seen from the figure 3a that APHRODITE (146-196 mm/seasonal) gives the closest precipitation estimates to the measured precipitation (275-584 mm/seasonal), followed by CRU TS 4.0.1 (155-194 mm/seasonal) and CMORPH_V1.0 (139-176 mm/seasonal) during summer season. In comparison with the measured precipitation, the reanalysis of ERA-Interim produces overestimation of precipitation for all temporal scales of annual, seasonal, month and daily. For all scales of temporal, spatial distribution of considered precipitation estimates tend to overestimate the precipitation towards the north and east of basin in comparison with the remaining sites, whereas measured precipitation towards the south of basin higher than the other sites. On the other hand, in comparison with measured precipitation, areas of high elevation seem to receive lower precipitation and higher precipitation receives over the areas of plain and coastal.

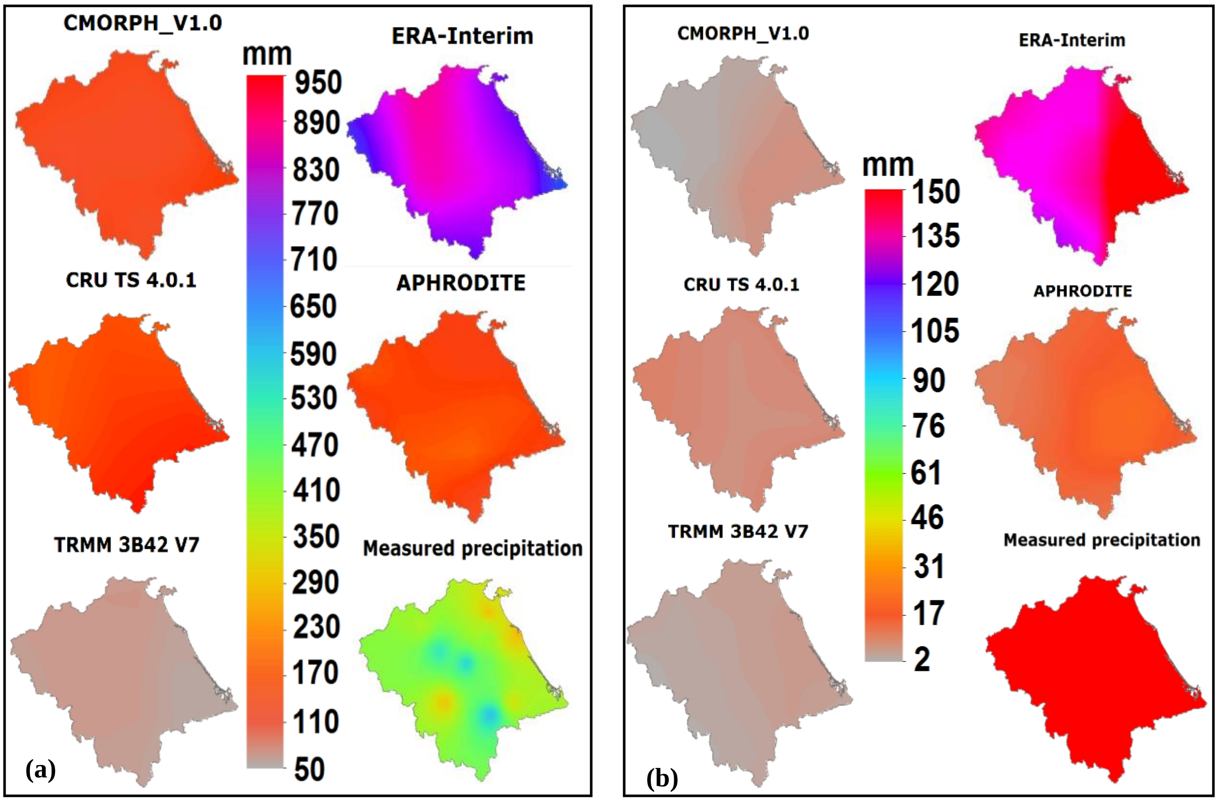

Figure 3. Spatial distribution of precipitation products for summer (a) and winter (b) seasons during 2007 - 2015 over VG-TB basin

3.2. Statistics of Continuous and Dichotomous

In this study, we examined the performance of the five precipitation estimations using the seven continuous statistics of RMSE, Bias, MAE, r, Ln(Nash), Nash and Mbias as summarized in Table 1. Figure 4 and Figure 5 represent the scatter plots of average monthly and daily precipitation, respectively, for precipitation products against the corresponding values from measured precipitation. Plus, the basin-averaged values (about 500 grid cells) of seven continuous statistical indices are shown in these figures. According to Figure 4a, APHRODITE obtained the best values for RMSE (200 mm/month), MAE (119 mm), r (0.45), Nash (0.4) and Ln(Nash) (0.96), whereas CRU TS 4.0.1 obtained the best values for Mbias (0.87) and Bias (-29.5 mm/month).

The values of Mbias (0.19), Bias (-194 mm/month) in the figure 4b and Mbias (0.27), Bias (-174.9 mm/month) in the figure 4e confirm that CMORPH_V1.0 and TRMM 3B42 V7 seriously underestimated precipitation levels over the study. Opposed to this, it is shown an extreme overestimation of ERA-Interim precipitation product with basin-averaged values of 1.7 and 170.2 (mm/month) for Mbias and Bias, respectively. At a scale of monthly, therefore, a suggestion of using the gridded precipitation products should be considered. It is possible to construct a long-term series of precipitation in the past using the APHRODITE product as the missing values of precipitation are often available in Vietnam in general and VG-TB basin in particular. Furthermore, as a suggestion that the monthly time series of APHRODITE should be tested to scale and adjust the TRMM 3B42 V7 and CMORPH_V1.0 with the purpose of 3-hour series time construction as follows:

$$Rain3hr_{C,T} = \frac{Rain_{month(A)}}{Rain_{month(C,T)}} Rain3hr_{C,T} \quad (1)$$

Where C denotes CMORPH_V1.0, T denotes TRMM 3B42 V7 and A refers APHRODITE.

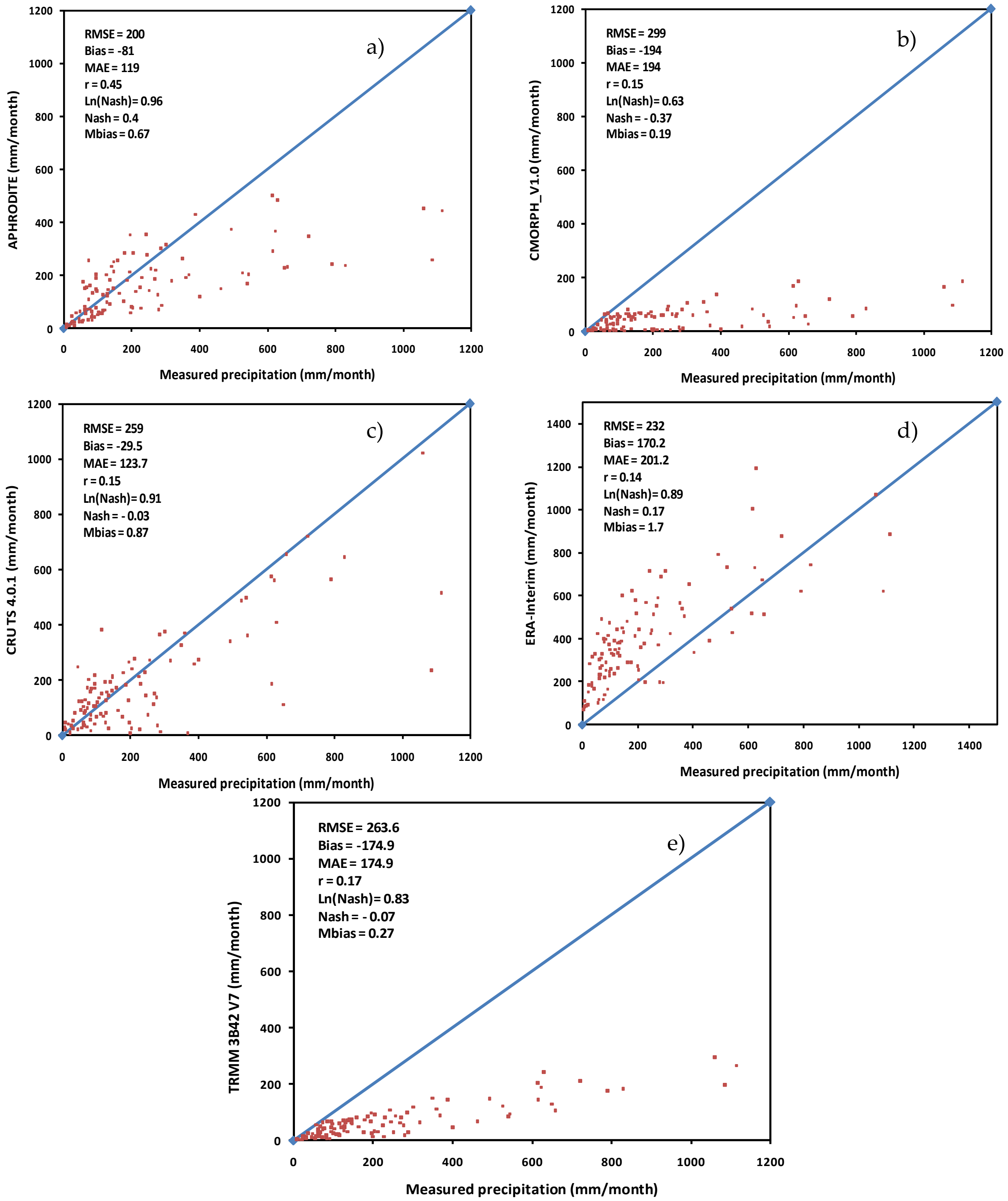

Figure 4. Scatter plots of monthly scale averaged values for (a) APHRODITE, (b) CMORPH_V1.0, (c) CRU TS 4.0.1, (d) ERA-Interim and (e) TRMM 3B42 V7 precipitation products against measured precipitation over VG-TB basin

At a daily scale, according to Figure 5a, APHRODITE obtained the best values for RMSE (15.3 mm), MAE (5.9 mm), Bias (-2.6 mm/day), Mbias (0.67) and Nash (0.38), whereas TRMM 3B42 V7 obtained the best values for r (0.04). The values of Mbias (0.19), Bias (-194 mm/month) in the figure 4b and Mbias (0.27), Bias (-174.9 mm/month) in the figure 4e confirm that CMORPH_V1.0 and TRMM 3B42 V7 seriously underestimated precipitation levels over the study. Opposed to this, it is shown an extreme overestimation of ERA-Interim precipitation product with basin-averaged values of 1.7 and 170.2 (mm/month) for Mbias and Bias, respectively.

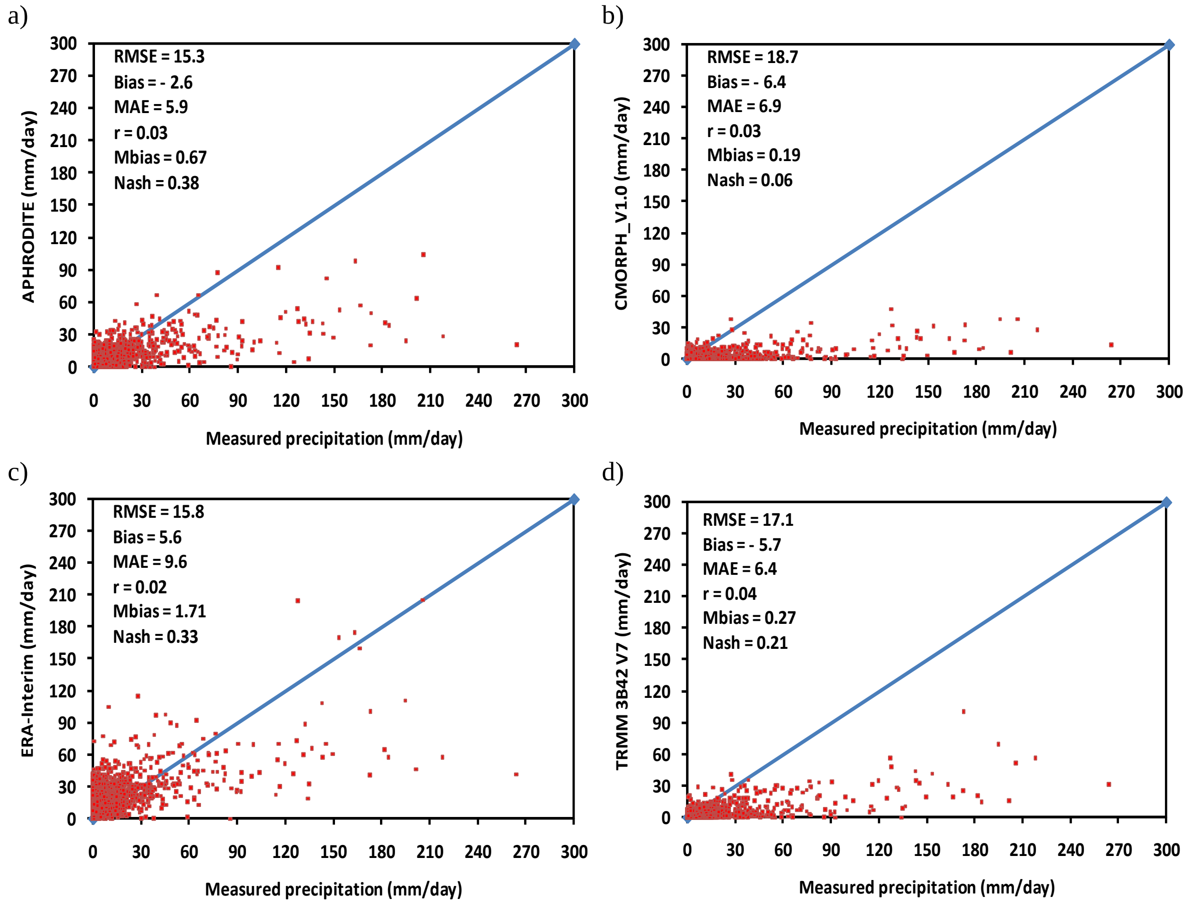

Figure 5. Scatter plots of daily scale averaged values for (a) APHRODITE, (b) CMORPH_V1.0, (c) ERA-Interim and (d) TRMM 3B42 V7 precipitation products against measured precipitation over VG-TB basin.

It is especially noteworthy that the performance of considered precipitation estimates well captures the minimum values of measured precipitation series with the Ln(Nash) index higher than 0.6. It can be seen the best product of APHRODITE [Ln(Nash),0.96], followed by CRU TS 4.0.1 [Ln(Nash), 0.91] and ERA-Interim [Ln(Nash),0.89]. On the other hand, APHRODITE is reliable very low rain events.

Table 4. Dichotomous statistics for considered precipitation estimates over VG-TB basin

Criteria	TRMM 3B42 V7	CMORPH_V1.0	APHRODITE	ERA-Interim
FAR	0.15	0.15	0.19	0.29
POD	0.68	0.55	0.83	0.98
CSI	0.6	0.5	0.69	0.7
POFD	0.24	0.19	0.39	0.78

Strongest **Weakest**

According to Table 4, TRMM 3B42 V7 and CMORPH_V1.0 get the best values for FAR (0.15), followed by APHRODITE (0.19) and ERA-Interim (0.29). More importantly, ERA-Interim presents the best values for POD (0.98) and CSI (0.7). Here, the value of POD, quite close to the perfect score of 1, indicates 98 percentages of the observed rain events are correctly reproduced, while over two-third of the "rain" events (observed and/or predicts) are correctly reproduced. For the POFD analysis, it is shown that CMORPH_V1.0 estimates are the best performance, followed by TRMM 3B42 V7. The POFD value of CMORPH_V1.0 (0.19) indicates that for 19 percentages of the observed "no rain" events the reproductions are incorrect. Contrary to this, in regard to FOD and CSI, CMORPH_V1.0 exhibits the worst performance.

Figure 6 shows a histogram of relative frequency classified and grouped the considered precipitation estimates and measured precipitation into 6 groups as mentioned in the section 2. It can be seen from this figure that ERA-Interim well captures rain events of heavy rain (50.0-100 mm/day). Contrary to this, rain events of zere rain reproduced by ERA-Interim mostly fail as displayed on figure 6. This figure refers that more or less 35% rain events are in category of light rain (0.6-6.0 mm/day).

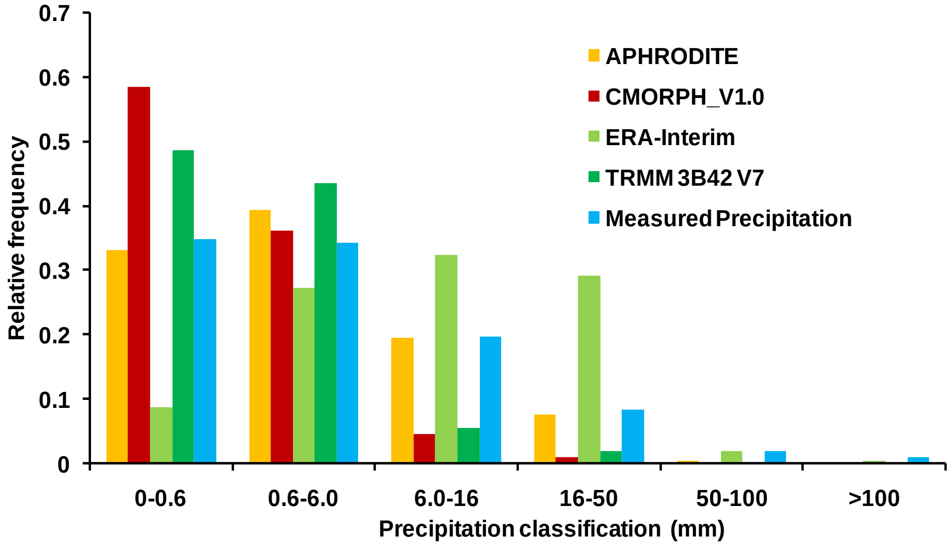

Figure 6. Histogram of relative frequency of daily considered precipitation estimates and measured precipitation during 2007 - 2015 over VG-TB basin

It is specially pay attention to the APHRODITE precipitation product that it better reproduces the rain events in all categories of rain events, except for extremely heavy rain above 100 mm/day. Plus, the cumulative distribution function of APHRODITE well matches the distribution of measured precipitation as shown in figure 7. Thus, it is deemed to be reliable and highly potential for hydrological and environmental applications at a temporal scale of daily without rain events of extremely heavy rain. On the other hand, all precipitation products completely fail to capture the rain events of extremely heavy rain. In comparison with measured precipitation, CMORHPH_V1.0 better reproduces the rain events with little overestimation of light rain (0.6-6 mm/day) than the others. For zero rain events (0-0.6 mm/day), TRMM 3B42 V7 gives the best performance.

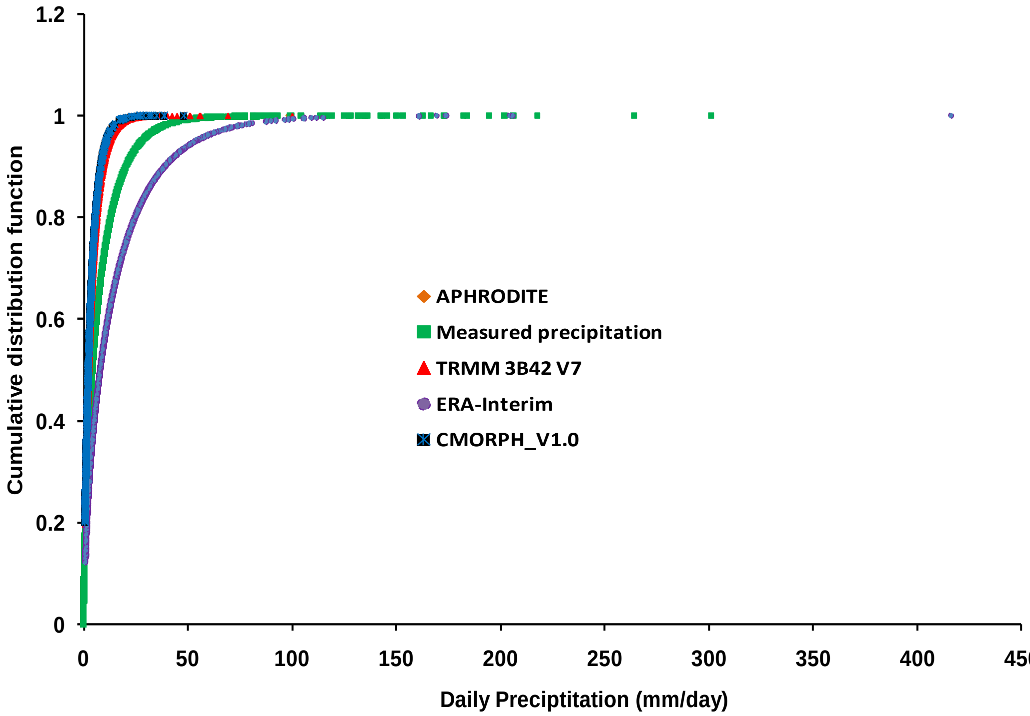

Figure 7. Cumulative distribution function of considered precipitation estimates in comparison with measured precipitation over VG-TB basin

4. Conclusion

This study evaluates multi-precipitation products for multi-time scales (daily, monthly, seasonal and annual) and spatial distribution for scale of basin with the complexity of climate regime. Overall, the cumulative distribution function of APHRODITE well matches the distribution of measured precipitation. The spatial distribution of considered precipitation estimates tend to overestimate the precipitation towards the north and east of basin in comparison with the remaining sites, while measured precipitation towards the south of basin higher than the other sites. So, there is a need to further investigation for mountain and coastal areas. ERA-Interim is able to capture the heavy precipitation events; remarkable overestimated the rainfall amount over the basin-averaged scale. TRMM 3B42 V7 is more accurate than CMORPH_V1.0. Author suggests that the monthly time series of APHRODITE should be tested to scale and adjust the TRMM 3B42 V7 and CMORPH_V1.0. CMORHPH_V1.0 better reproduces the rain events with little overestimation of light rain (0.6-6 mm/day) than the other precipitation products. For zero rain events (0-0.6 mm/day), TRMM 3B42 V7 shows the best performance of zero rain events (0-0.6 mm/day) in comparison with other precipitation products. Specially, there is a need to add ERA-Interim for a correction the daily precipitation.

5. Conflict of Interest

The author declares no conflict of interest.

6. References

[1] Yatagai, Akiyo, Kenji Kamiguchi, Osamu Arakawa, Atsushi Hamada, Natsuko Yasutomi, and Akio Kitoh. "APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges." Bulletin of the American Meteorological Society 93, no. 9 (2012): 1401-1415. doi:10.1175/bams-d-11-00122.1.

[2] Ushio-Fukai, Masuko. "Compartmentalization of redox signaling through NADPH oxidase–derived ROS." Antioxidants & redox signaling 11, no. 6 (2009): 1289-1299. doi:10.1089/ars.2008.2333.

[3] Peterson, Thomas C., and Russell S. Vose. "An overview of the Global Historical Climatology Network temperature database." Bulletin of the American Meteorological Society 78, no. 12 (1997): 2837-2850. doi:10.1175/1520-0477(1997)078<2837:AOOTGH>2.0.CO;2.

[4] Rodell, Matthew, P. R. Houser, U. E. A. Jambor, J. Gottschalck, K. Mitchell, C-J. Meng, K. Arsenault et al. "The global land data assimilation system." Bulletin of the American Meteorological Society 85, no. 3 (2004): 381-394. doi:10.1175/BAMS-85-3-381.

[5] Joyce, Robert J., John E. Janowiak, Phillip A. Arkin, and Pingping Xie. "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution." Journal of Hydrometeorology 5, no. 3 (2004): 487-503. doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

[6] Adler, Robert F., George J. Huffman, Alfred Chang, Ralph Ferraro, Ping-Ping Xie, John Janowiak, Bruno Rudolf et al. "The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present)." Journal of hydrometeorology 4, no. 6 (2003): 1147-1167. doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

[7] Harris, I. P. D. J., Philip D. Jones, Timothy J. Osborn, and David H. Lister. "Updated high‐resolution grids of monthly climatic observations‐the CRU TS3. 10 Dataset." International journal of climatology 34, no. 3 (2014): 623-642. doi:10.1002/joc.3711.

[8] Drobot, Sheldon, James Maslanik, Ute Christina Herzfeld, Charles Fowler, and Wanli Wu. "Uncertainty in temperature and precipitation datasets over terrestrial regions of the Western Arctic." Earth Interactions 10, no. 23 (2006): 1-17. doi:10.1175/ei191.1.

[9] Nguyen, Tien Thanh. "Improved Downscaling of Meteorological Data for Hydrological Modeling in the Tropics Under Climate Change." PhD diss., Technische Universität Carolo-Wilhelmina zu Braunschweig, 2016. https://nbn-resolving.org/urn:nbn:de:gbv:084-16072812348.

[10] Xie, Pingping, Mingyue Chen, Song Yang, Akiyo Yatagai, Tadahiro Hayasaka, Yoshihiro Fukushima, and Changming Liu. "A gauge-based analysis of daily precipitation over East Asia." Journal of Hydrometeorology 8, no. 3 (2007): 607-626. doi:10.1175/jhm583.1.

[11] Nguyen, Tien Thanh. "Fitting a Probability Distribution to Extreme Precipitation for a Limited Mountain Area in Vietnam." Journal of Geoscience and Environment Protection, 2017. 5(05): p. 92. doi:10.4236/gep.2017.55007.

[12] Nguyen, Tien Thanh and Luca Dutto Aldo Remo. "Projected changes of precipitation idf curves for short duration under climate change in central Vietnam." Hydrology 5, no. 3 (2018): 33. doi:10.3390/hydrology5030033.

[13] WMO-Nr.8. "Guide to Meteorological Instruments and Methods of Observation." World Meteorological Organisation: Geneva, Switzerland, 2008. Available online: http://library.wmo.int/opac/index.php?lvl=notice_display&id=12407#.WDVYPdXyuM8 (accessed on 03 September 2018).

[14] WMO-Nr.100. "Guide to Climatological Practices." 2nd ed. 1983. Available online: https://library.wmo.int/pmb_ged/wmo_100_en.pdf (accessed on 03 September 2018).

[15] Fekete, Balázs M., Charles J. Vörösmarty, John O. Roads, and Cort J. Willmott. "Uncertainties in precipitation and their impacts on runoff estimates." Journal of Climate 17, no. 2 (2004): 294-304. doi:10.1175/1520-0442(2004)017<0294:uipati>2.0.co;2.

[16] Bosilovich, Michael G., Junye Chen, Franklin R. Robertson, and Robert F. Adler. "Evaluation of global precipitation in reanalyses." Journal of applied meteorology and climatology 47, no. 9 (2008): 2279-2299. doi:10.1175/2008jamc1921.1.

[17] Bukovsky, Melissa S., and David J. Karoly. "A brief evaluation of precipitation from the North American Regional Reanalysis." Journal of Hydrometeorology 8, no. 4 (2007): 837-846. doi:10.1175/jhm595.1.

[18] Goudenhoofdt, E., and L. Delobbe. "Evaluation of radar-gauge merging methods for quantitative precipitation estimates." Hydrology and Earth System Sciences 13, no. 2 (2009): 195-203. doi:10.5194/hessd-5-2975-2008.

[19] Peña-Arancibia, Jorge L., Albert IJM van Dijk, Luigi J. Renzullo, and Mark Mulligan. "Evaluation of precipitation estimation accuracy in reanalyses, satellite products, and an ensemble method for regions in Australia and South and East Asia." Journal of Hydrometeorology 14, no. 4 (2013): 1323-1333. doi:10.1175/jhm-d-12-0132.1.

[20] Pfeifroth, Uwe, Richard Mueller, and Bodo Ahrens. "Evaluation of satellite-based and reanalysis precipitation data in the tropical Pacific." Journal of Applied Meteorology and Climatology 52, no. 3 (2013): 634-644. doi:10.1175/jamc-d-12-049.1.

[21] Tramblay, Y., Vera Thiemig, Alain Dezetter, and Labhoucine Hanich. “Evaluation of satellite-based rainfall products for hydrological modelling in Morocco.” Hydrological Sciences Journal, 2016. 61(14): p. 2509-2519. doi:10.1080/02626667.2016.1154149.

[22] Kenabatho, P., B. Parida, and D. Moalafhi. “Evaluation of satellite and simulated rainfall products for hydrological applications in the Notwane Catchment, Botswana.” Physics and Chemistry of the Earth, Parts A/B/C, 2017. 100: p. 19-30. doi:10.1016/j.pce.2017.02.009.

[23] Ayehu, G.T., Tsegaye Tadesse, Berhaan Gessesse, and Tufa Dinku. “Validation of new satellite rainfall products over the Upper Blue Nile Basin, Ethiopia.” Atmospheric Measurement Techniques, 2018. 11(4): p. 1921-1936. doi:10.5194/amt-11-1921-2018.

[24] MONRE. “QUY HOẠCH MẠNG LƯỚI QUAN TRẮC TÀI NGUYÊN VÀ MÔI TRƯỜNG QUỐC GIA GIAI ĐOẠN 2016 - 2025, TẦM NHÌN ĐẾN NĂM 2030” in Vietnames, 2015. Available online: http://datafilesbk.chinhphu.vn/file-remote-v2/DownloadServlet?filePath=vbpq/2016/01/90.signed.pdf (accessed on 03 September 2018).

[25] Huffman, George J., and David T. Bolvin. "TRMM and other data precipitation data set documentation." NASA, Greenbelt, USA 28 (2013). Available online: https://pmm.nasa.gov/sites/default/files/document_files/3B42_3B43_doc_V7.pdf (accessed on 29 August 2018).

[26] Ferraro, Ralph R. "Special sensor microwave imager derived global rainfall estimates for climatological applications." Journal of Geophysical Research: Atmospheres 102, no. D14 (1997): 16715-16735. doi:10.1029/97JD01210.

[27] Ferraro, Ralph R., Fuzhong Weng, Norman C. Grody, and Limin Zhao. "Precipitation characteristics over land from the NOAA‐15 AMSU sensor." Geophysical Research Letters 27, no. 17 (2000): 2669-2672. doi:10.1029/2000GL011665.

[28] Kummerow, Christian, Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G. Petty, Dong-Bin Shin, and T. T. Wilheit. "The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors." Journal of Applied Meteorology 40, no. 11 (2001): 1801-1820. doi:10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2.

[29] Dee, Dick P., S. M. Uppala, A. J. Simmons, Paul Berrisford, P. Poli, S. Kobayashi, U. Andrae et al. "The ERA‐Interim reanalysis: Configuration and performance of the data assimilation system." Quarterly Journal of the royal meteorological society 137, no. 656 (2011): 553-597. doi:10.1002/qj.828.

[30] Berrisford, Paul, D. P. K. F. Dee, Keith Fielding, Manuel Fuentes, P. Kallberg, Shinya Kobayashi, and Sakari Uppala. "The ERA-interim archive." ERA report series 1 (2009): 1-16. Available online: https://www.ecmwf.int/en/elibrary/8174-era-interim-archive-version-20 (accessed on 30 August 2018).