# *Evaluation of Satellite Precipitation from Google Earth Engine in Tonle Sap Basin, Cambodia*

**Phanit Mab[1], Steven Ly[2], Chuphan Chompuchan[1], Ekasit Kositsakulchai[1,3,a,*]**

**Abstract** Precipitation is important to life on Earth. It is a predominant process in the global hydrologic cycle and is an indispensable component of water balance analysis. However, in some area like the Tonle Sap basin in Cambodia, the information on precipitation is deficient and sometimes difficult to access. In this case, satellite remote sensing coupled with GIS techniques have been applied and considered as a powerful and effective tool in handing precipitation analysis tasks. Recently, the Google Earth Engine (GEE) platform provides satellite datasets and collection of the tool for analysis of data using JavaScript without downloading huge data from the Internet. In this study, we aimed to evaluate the application of GEE platform for retrieving and analyzing precipitation data of the Tropical Rainfall Measuring Mission (TRMM) in Tonle Sap basin (TLS). The methods included: (1) to collect the satellite precipitation data (3B43V7) by manual download and by retrieving them from GEE platform; (2) to analyze monthly precipitation over the study area by GIS analysis functions and by JavaScript on the GEE platform, data in 2010 was sampling as a case study; and (3) to compare results from both GIS and GEE with observation data from ground stations. The results showed the good correlations between the precipitations from manual download and those from a GEE platform, with R greater than 0.9. In short, the application of GEE platform is very effective; it provides a comprehensive tool for managing time-consuming tasks, like precipitation data collection and analysis, and results in reliable outputs.

**Keywords** *Satellite precipitation, Mekong River, Google Earth Engine*

[1]Research Laboratory for Intelligent system in Agro-hydrological Monitoring and Management (INAMM)
Faculty of Engineering
at Kamphaeng Sean, Kasetsart University,
Nakhon Pathom, 73140 Thailand.
[2]Department of Civil and Earth Resources Engineering, Kyoto University, Japan.
[3]Water Engineering and Management Program,
Asian Institute of Technology,
Pathumthani, 12120 Thailand.

[a]ekasit.k@ku.th (corresponding author)

## Introduction

Precipitation is a critical variable in the global hydrologic cycle, and it influences our daily lives (drought, floods, agricultural, irrigation, outdoor activities, etc.)[1]. However, the accessibility of precipitation data in the Tonle Sap Lake region is one of limitations to conduct a comprehensive hydrological analysis.

In this case, the techniques of satellite remote sensing have been widely used and been considered as a powerful and effective tool in perceiving precipitation. However, a massive datasets have to be downloaded; satellite image processing with geographic information system (GIS) analysis function is the prerequisite before retrieval ofrainfall information.

Recently, the Google Earth Engine (GEE) leverages cloud computing services to provide analysis capabilities on over 40 years of Landsat data[2], and others satellites. As a remote sensing platform, its ability to analyze global data rapidly lends itself to being an useful tool on data visualization [3]. Additionally, dataset is processing of geospatial datasets an online for rapid visualization of complex spatial analyses using the Javascript Application Programming Interface (API). This API allows us to develop a code in order to get datasets of publicly available remotely sensed imagery and other data.

This study aimed to apply the GEE platform for automatically retrieving and analyzing precipitation data of the Tropical Rainfall Measuring Mission (TRMM) in Tonle Sap Lake basin (TLS), Cambodia. The methods included: (1) to collect the monthly precipitation data (3B43V7) by manually download and to retrieve from GEE platform; and (2) to compare a selected precipitation data of whole TLS basin in 2010, derived by GIS analysis functions and by JavaScript on a GEE platform, and (3) to compare both GIS and GEE with observation data from ground stations.

## Material and Method

### *A. Study Area*

The Tonle Sap Lakeis the largest permanent freshwater lake, locating in Cambodia at the Lower Mekong River basin [4]. During the dry season, the lake is about 120 km long and 35 km widewith an area of about 2500 $km^2$. A bathymetrical survey of the lake proper, conducted between 1997 and 1999, revealed a relatively flat bottom, with a maximum depth of about 3.3 m[5].During the flood period the Lakeexpands to 250 km long and 100 km wide with an area of about 17,500 $km^2$, and the depth reaches 8-10 m. The floodplain surrounding the lake extends 20-40 km and is dominated by seasonally inundated forest and rice field [5, 6]. The Tonle Sap basin, extends over 44% of Cambodia's total area 80,000 $km^2$. 32% of Cambodia's total populationdepend on this lake in living [7].

There arefiveprovinces bordering Tonle Sap Lake namely: Siem Reap, Battambang, Pusat, Kampong Cham, and Kampong Thom. Tonle Sap Basin consists of 11 sub-basins: namely (1) Stung Sreng, (2) Stung Chikreng, (3) Tonle Sap, (4) Stung Pursat, (5) Stung Dauntri, (6) Stung Boribo, (7) Stung Sangker, (8) Stung Monkong Borey, (9) Stung Staung, (10) Stung Sen, (11) Stung Chinit and (12) Stung Sieamreap (Fig. 1).

Tonle Sap basin climate influences from the tropical monsoon seasons. Dry season runs from December to April and rainy season comes when the winds shift into the southwestmonsoonfrom May to November. The monsoon returns south during August and October when the rainfall is usually heavier, with the highestrainfall in October.

*B. Satellite Precipitation-TRMM 3B43V7*

Many studies have been conducted on detecting diurnal cycles of precipitation over different parts of the world usingsatellite data, especially the Tropical Rainfall Measuring Mission (TRMM)[8].

Recently, data 3B43 is monthly executed to produce the precipitation rate field (3B43). These were combining the 3-hourly merged high-quality/IR estimates (3B42) with the monthly accumulated Global Precipitation Climatology Centre (GPCC) rain gauge analysis. Data are available from 1998 to presentat https://pmm.nasa.gov/data-access/downloads/trmm.

*C. Precipitation Processing by GEE Platform*

Google Earth Engine (GEE)platform facilitates a fast analysis by using Google's cloud-computing infrastructure(https://earthengine.google.org/).The pre-processed monthly data of Tropical Rainfall Measuring Mission data, or TRMM3B43V7, available through GEE was used to assess precipitation data across the study area. The spatial resolution of the above datasets are all equal to $0.25° \times 0.25°$ with monthly gridded rainfall data [9, 10].

This process is using JavaScript coding in GEE platform screen. TRMM 3B43V7 datasets could be specified location of boundary of Tonle Sap basins in coding. At the end, the amount of precipitation data had retrieved each sub-basins in CSV which can open with MS Excel to interpret data in number.

*D. Thiessen Polygon Method*

The Thiessen polygon methodallows for areal weighting of rainfall from each gauge. $A_i$ is polygon area, $P_i$ is average precipitation and A is total area [11].

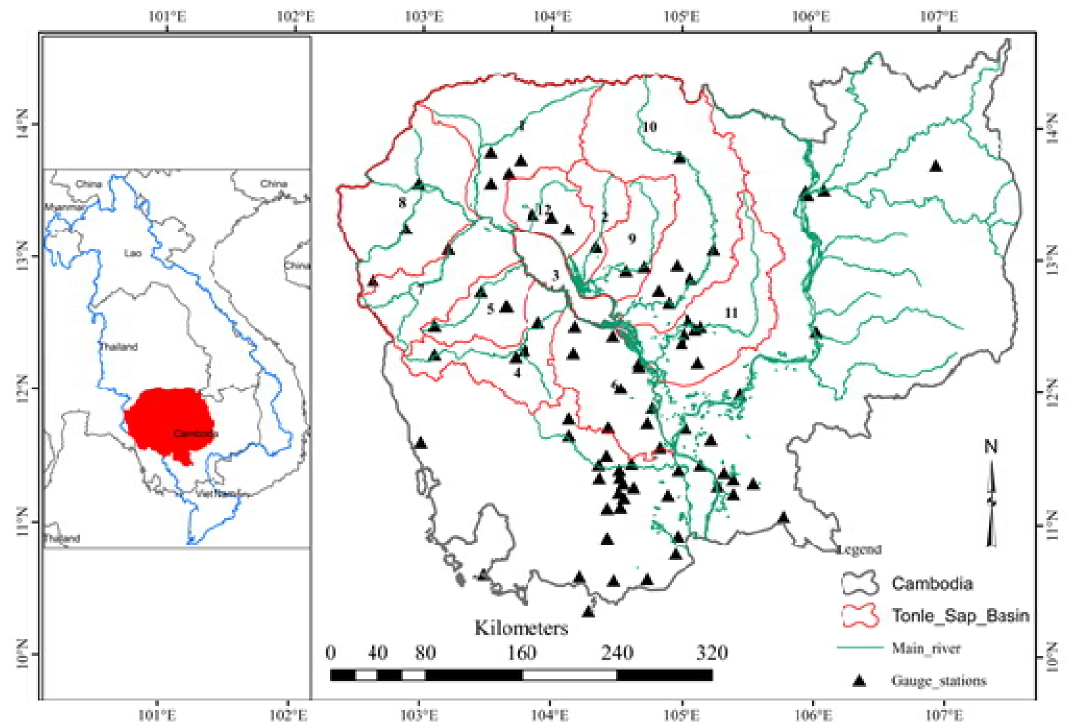



**Fig. 1.** Boundary of Tonle Sap Basin and location raingauge stations

*E. Statistical Indicators*

Correlation coefficient (r) is measured on a scale between -1 and +1 to determine the extent to which two sets of paired values are related in a linear fashion. The number closer to -1 and +1 are -1 is a perfect negative linear correlation, whereas the number closer to +1 is a perfect positive linear correlation, and 0 is no correlation.

$$r = \frac{\sum_{j=1}^{N}(X_j - \bar{X})(Y_j - \bar{Y})}{\sqrt{\sum_{j=1}^{N}(X_j - \bar{X})^2 \sum_{j=1}^{N}(Y_j - \bar{Y})^2}} \quad (2)$$

Root mean square error (RMSE) is the square root of the mean of the summation of squared differences between two sets of values where there are n number of paired values x and y. This statistic provides an absolute (neither positive nor negative) value of differences between two sets of values. A smaller value signifies less error.

$$RMSE = \sqrt{\frac{1}{N}\sum_{j-1}^{N}(x_j - y_j)^2} \quad (3)$$

*F. Methodology*

In this study, 94 rainfall stations from Mekong River Commission (MRC) and Cambodia's Ministry of WaterResources and Meteorology (MoWRAM) were selected which covered almost the entire of Tonle Sap basin.

To achieve the research objectives, data processing methods were divided into three main parts (Fig. 2), including manual data acquisition and GIS processing, GEE-based data processing and gauge observation using Thiessen Polygon method.

The first part is to download monthly precipitation in 2010 (TRMM 3B43v7) in raster format (HDF), and use GIS analysis functions to interpret data, change coordinate system to the same area. Using the function "Zonal statistic as table" in order to get the amount of precipitation with "shape file of boundary".

The second part is to develop the JavaScript code on the GEE platform (https://code.earthengine.google.com).The code included: importing the TRMM image dataset; specifying the boundary of study area by uploading shape file into GEE data script; extracting the mean monthly precipitation of the basin by writing code on script (see example in Fig. 3). After having execute the code, the precipitation data were retrieved on sub-basins automatically.

Lastly, the areal precipitation from gauge station were estimated by Thiessen Polygon method. Precipitation of basin was summarized using "Zonal statistic as table" with boundary of basin. The precipitation data retrieved from the first and second methods were finally compared with the observed data from gauge stations for the entire basin and each sub-basin. Statistical indicators, correlation coefficient (r) and RMSE, were estimated.

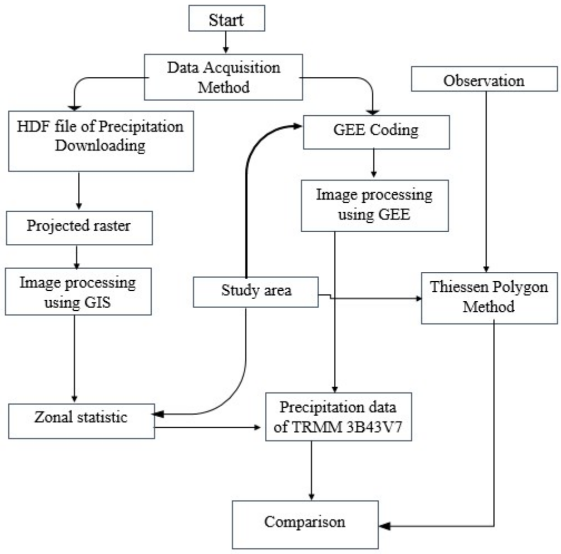



**Fig. 2.** Research framework

**Results and discussions**

*A. Monthly Precipitaion Depth*

Fig. 4 shows monthly precipitation of the Tonle Sap Basin in 2010, derived from 3 methods: GIS analysis of downloaded TRMM data (GIS), GEE processing (GEE) and observed rainfall with Thiessen Polygon estimation (OBS).

*1) TRMM Precipitation from GIS vs. GEE Processings*

The amount of precipitation from TRMM 3B43V7 using both GIS analysis of downloaded data (GIS) and Google Earth Engine processing (GEE) methods were well matched with almost identical rainfall depths from January to December on 2010. In dry season from November to April, rainfall varied from 8 to 70 mm, while in wet season from May to October, rainfall varied from 100 to 300 mm. The highest precipitation can be observed in October around 300 mm. GEE provided an almost similar the amount of rainfall compared to manually download rainfall. There are somewhat difference between 0.005 to 1.8 mm due to round-off errors during data processing.

*2) Precipitation from GEE Processings vs. Observation*

From Fig. 4, TRMM precipitation from GEE processing and basin rainfall estimation from gauge station (OBS) were well captured the amount of monthly rainfall. By comparing GEE rainfall with rain gauge data, we can observed an over-estimated rainfall by GEE mostly during dry period from January to June (4 to 20 mm), in October and in December. Rainfall by GEE were under-estimated during wet period from July to September 30 to 49 mm. However, these values were observed as less error between rainfall from GEE and observation. Fig. 5 showed the spatial data on the monthly average rainfall in millimeters from the GEE platform.

*3) Precipitation from GIS Processing vs. Observation*

Based on result shown in Fig.4, the differences of rainfall depths between GIS processing (GIS) and the gauged observation (OBS) are comparable to those between GEE and OBS. The identical months of over- or under-estimation were observed. The GIS rainfall were over-estimated from the lowest of 4 mm on December to the highest of 20 mm on June.

## B. Statistical Indicators

Statistical indicators of the whole basin were shown in Table I, the amount of precipitations from manual download and those from a GEE platform provided an almost similar monthly rainfall depth with r greater than 0.99 whole basin. Finally, both methods GEE and GIS provided a good correlation with the gauge which showed the r value greater than 0.97 and the highest RMSE of basin about 16 mm.

Statistical indicators of each sub-basins were presented in Table II, the indicators demonstrated the good correlations of rainfall between GIS *vs.* GEE, GEE *vs.* observation, and GIS *vs.* observation. The correlation coefficient, r, are 0.99, 0.81 and 0.80 respectively.

**Table I.** Statistical Indicators of Basin

| Method | *r* | *RMSE* |
|---|---|---|
| GIS vs. GEE | 0.999 | 0.70 |
| GEE vs. OBS | 0.975 | 16.40 |
| GIS vs. OBS | 0.977 | 15.88 |

**Table II.** Statistical Indicator of each Sub-basins

| Sub basin | GIS *vs.* GEE | | GIS *vs.* OBS | | GEE *vs.* OBS | |
|---|---|---|---|---|---|---|
| | *r* | *RMSE* | *r* | *RMSE* | *r* | *RMSE* |
| 1 | 0.9898 | 1.20 | 0.8713 | 23.56 | 0.8818 | 23.97 |
| 2 | 0.9997 | 1.13 | 0.8719 | 27.21 | 0.8259 | 27.73 |
| 3 | 0.9998 | 1.73 | 0.9611 | 17.72 | 0.9347 | 17.94 |
| 4 | 0.9998 | 2.15 | 0.9723 | 24.81 | 0.9424 | 25.80 |
| 5 | 0.9905 | 1.39 | 0.8437 | 23.49 | 0.8824 | 23.62 |
| 6 | 0.9996 | 2.36 | 0.8089 | 20.73 | 0.8984 | 21.46 |
| 7 | 0.9991 | 2.36 | 0.9032 | 39.27 | 0.8010 | 38.77 |
| 8 | 0.9997 | 7.26 | 0.8864 | 29.70 | 0.8725 | 26.18 |
| 9 | 0.9997 | 0.86 | 0.9392 | 19.51 | 0.9697 | 19.82 |
| 10 | 0.9994 | 1.30 | 0.9393 | 19.99 | 0.9638 | 20.14 |
| 11 | 0.9997 | 1.00 | 0.8326 | 24.96 | 0.8757 | 24.64 |
| 12 | 0.9997 | 6.59 | 0.8872 | 28.41 | 0.8897 | 27.94 |

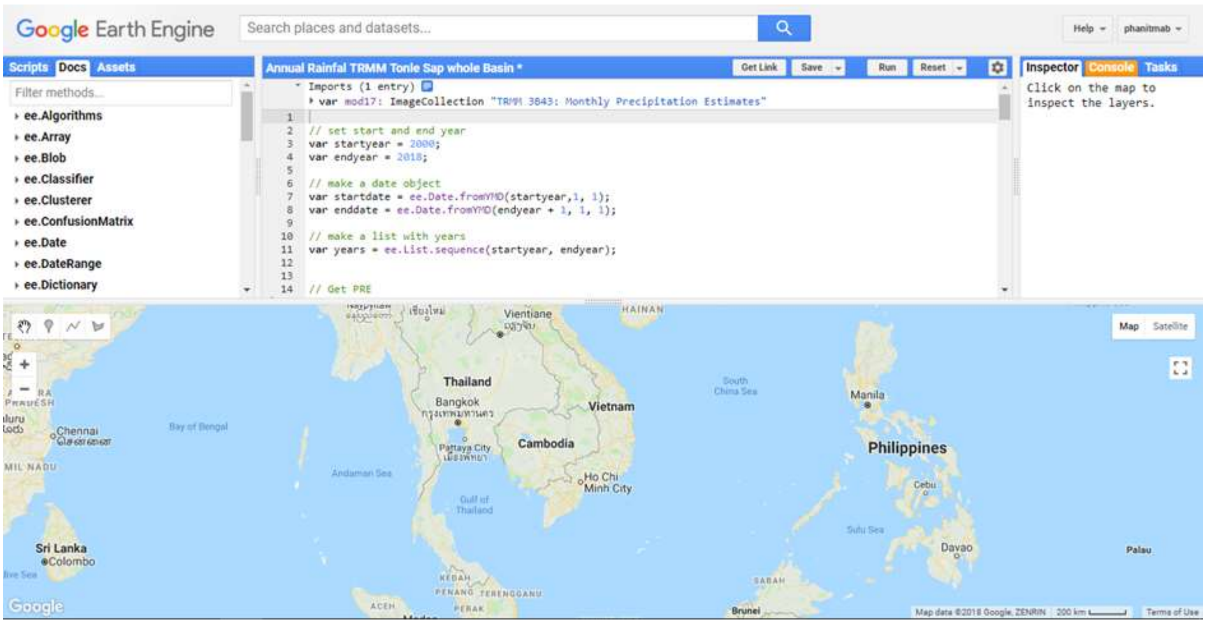
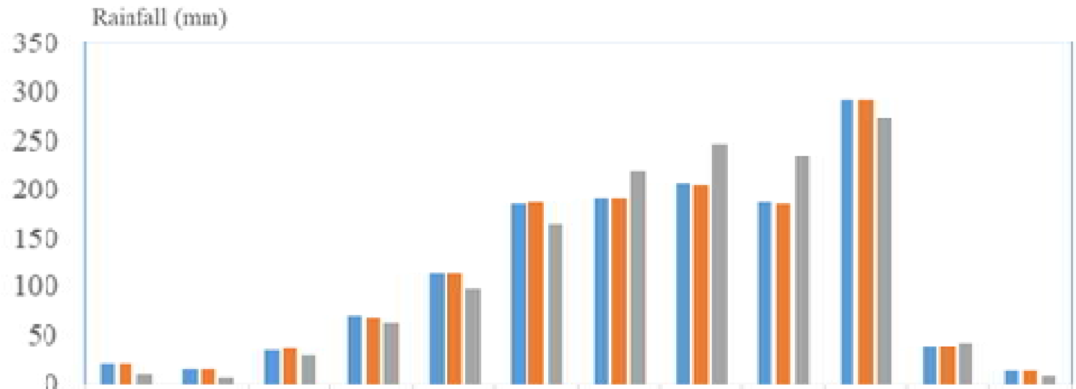



**Fig. 3.** Example of GEE platform interface for developer




| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GIS | 20.5 | 15.8 | 33.8 | 68.6 | 112.6 | 185.0 | 189.5 | 205.6 | 186.3 | 291.2 | 37.5 | 12.8 |
| GEE | 20.9 | 15.8 | 35.3 | 67.7 | 112.3 | 186.5 | 188.6 | 203.8 | 185.0 | 291.6 | 38.2 | 13.0 |
| Observe | 11.0 | 7.0 | 28.7 | 61.6 | 97.2 | 164.2 | 217.2 | 246.3 | 234.9 | 273.6 | 42.7 | 8.1 |

GIS GEE Observe

**Fig. 4.** Montly rainfall in 2010 estimated by three methods: GIS analysis of downloaded data (GIS), GEE processing (GEE) and observed rainfall with Thiessen Polygon estimation (OBS)

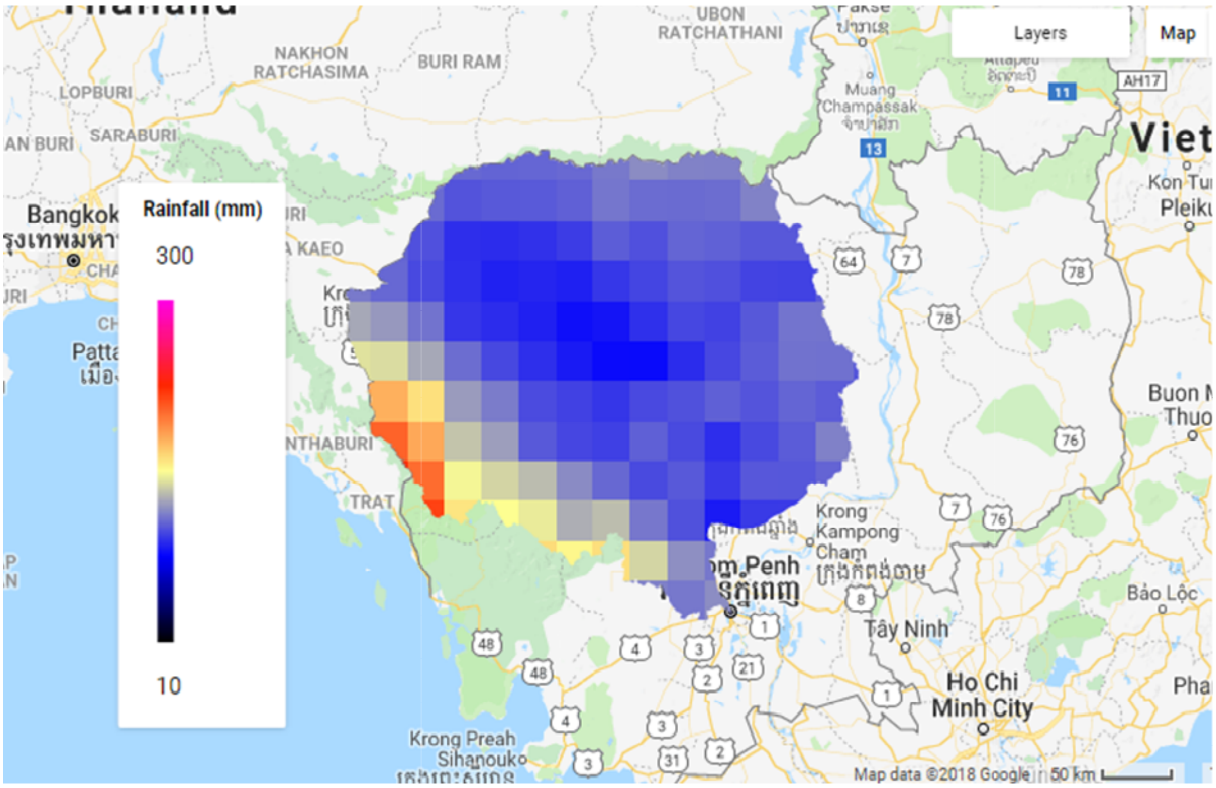



**Fig. 5.** Spatial distribution of monthly rainfall overlayed on Google Map processed by GEE platform

View publication stats



## Conclusions

Both precipitations from manually download and those froma GEE platform provided an almost similar rainfall depth for each sub-basins and the entire basin. However, the amount of precipitation of each sub-basins had some error because of uncertainty and location of the gauges in basins. TRMM 3B43V7 dataset provided high correlation when compared with the gauge stations. In short, the GEE platform is an effective tool which provides a comprehensive for managing time-consuming tasks, namely precipitation data collection and analysis. It is a new concept of remote sensing platform on how to get satellite datasets easily and quickly with results in reliable outputs.

## Acknowledgment

This paper was supported by the Kasetsart University Scholarships for ASEAN for Commemoration of the 60th Birthday Anniversary of Professor Dr. Her Royal Highness Princess Chulabhorn Mahidol. Moreover, the authors would also like to express their gratitude to the Google Earth Engine (GEE) developers for providing a registration during this research on retrieval datasets.

## References

[1] G. Langella, A. Basile, A. Bonfante, and F. Terribile, "High-resolution space–time rainfall analysis using integrated ANN inference systems," Journal of hydrology, vol. 387, no. 3-4, pp. 328-342, 2010.
[2] J. Dong, X. Xiao, M. A. Menarguez, G. Zhang, Y. Qin, D. Thau, C. Biradar, and B. Moore III, "Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine," Remote sensing of environment, vol. 185, pp. 142-154, 2016.
[3] N. N. Patel, E. Angiuli, P. Gamba, A. Gaughan, G. Lisini, F. R. Stevens, A. J. Tatem, and G. Trianni, "Multitemporal settlement and population mapping from Landsat using Google Earth Engine," International Journal of Applied Earth Observation and Geoinformation, vol. 35, pp. 199-208, 2015.
[4] I. Campbell, S. Say, and J. Beardall, "Tonle Sap Lake, the heart of the lower Mekong," in The Mekong: Elsevier, 2009, pp. 251-272.
[5] I. C. Campbell, C. Poole, W. Giesen, and J. Valbo-Jorgensen, "Species diversity and ecology of Tonle Sap Great Lake, Cambodia," Aquatic Sciences, vol. 68, no. 3, pp. 355-373, 2006.
[6] W. J. Junk, M. Brown, I. C. Campbell, M. Finlayson, B. Gopal, L. Ramberg, and B. G. Warner, "The comparative biodiversity of seven globally important wetlands: a synthesis," Aquatic Sciences, vol. 68, no. 3, pp. 400-414, 2006.
[7] ADB, "The Tonle Sap Basin Strategy," 2005, p. 9.
[8] M. S. K. Kiany, R. C. Balling Jr, R. S. Cerveny, and D. S. Krahenbuhl, "Diurnal variations in seasonal precipitation in Iran from TRMM measurements," Advances in Space Research, 2018.
[9] Z. Duan and W. Bastiaanssen, "First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure," Remote Sensing of Environment, vol. 131, pp. 1-13, 2013.
[10] G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker, "The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales," Journal of hydrometeorology, vol. 8, no. 1, pp. 38-55, 2007.
[11] P. B. Bedient, W. C. Huber, and B. E. Vieux, Hydrology and floodplain analysis. 2008.