Rainfall variability in Malay Peninsula region of Southeast Asia using gridded data

Vishal Singh[1] *and Xiaosheng* Qin[1*]

[1]School of Civil and Environmental Engineering, Nanyang Technological University, Singapore 639798

Abstract Southeast Asia is recognized as a climate-change vulnerable region as it has been significantly affected by many extreme events in the past. This study carried out a rainfall analysis over the Malay Peninsula region of Southeast Asia utilizing historical (1981-2007) gridded rainfall datasets (0.5°×0.5°). The rainfall variability was analyzed in an intra-decadal time series duration. The uncertainty involved in all datasets was also checked based on the comparison of multiple global rainfall datasets. Rainfall gap filling analysis was conducted for producing more accurate rainfall time series after testing multiple mathematical functions. Frequency-based rainfall extreme indices such as Dry Days and Wet days are generated to assess the rainfall variability over the study area. Our results revealed a notable variation existed in the rainfalls over Malay Peninsula as per the long historical duration (1981-2007).

1 Introduction

Southeast Asia has corresponded with two major types of monsoon systems, i.e. South Asia Summer Monsoon (SASM) and Boreal Winter Season (BWS) or Asian Winter Monsoon [1]. These two monsoon systems are accountable for the wet and dry seasons in the SEA [1]. Malay Peninsula is mostly influenced by BWS (North-East) which brings heavy rainfalls mostly during November to March [2, 3]. The resolution and time step of rainfall data imply the accuracy of any rainfall-based analysis outcomes. The missing values in rainfall time series is common but long data gaps could be critical since these gaps in rainfall time series can produce biased results [4, 5]. To reduce the uncertainty in the filled rainfall time series, the main prominence should be maintaining rainfall frequency, extremity and their patterns [6]. The objective of this study is to analyse rainfall variabilities over the Malaya Peninsula region in a relatively long-term (1981-2007) duration using gridded rainfall dataset at a resolution of 0.50°×0.50°. The gridded rainfall dataset is corrected using various gap-filling methods. The seasonality and rainfall patters are also explored.

2 Study area and data

The selected study area covers the upper and lower bounds within Malay Peninsula region, namely between -1°S to 7°N (latitude) and 100°E to 104°E (longitude). The seasonality of SEA region including Malay Peninsula and their complexities was discussed by Mandapaka et al. [7]. The daily high resolution observed gridded rainfall

*Corresponding author: xsqin@ntu.edu.sg

dataset (SA-OBS), generated under the Southeast Asian Climate Assessment and Dataset (SACA&D) project [8], was adopted in this study. The SACA&D used more than 1393 rainfall gauges in the preparation of SA-OBS gridded rainfall datasets. In this study, we utilized the SA-OBS data from 1981 to 2007 produced at 0.50°×0.50° grid scale (see Fig. 1).

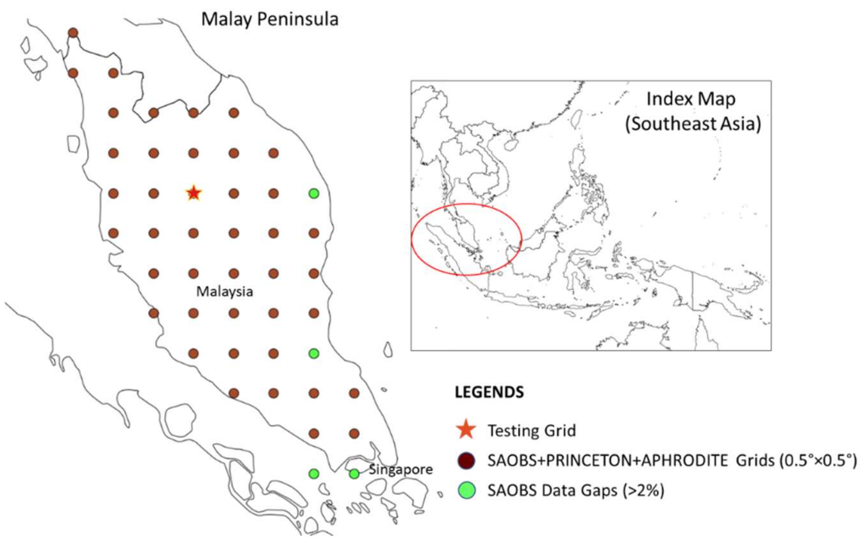

Fig. 1. The study Area map of Southeast Asia covering Malay Peninsula region. The rainfall grids used in the study are highlighted within the Malay Peninsula region.

The Terrestrial Hydrology Research Group Princeton University provides the gridded daily rainfall dataset (1981-2007) at 0.50°×0.50° grid scale for the whole world (*http://hydrology.princeton.edu/data.pgf.php*) and it was also utilized in our study [9]. This dataset has been utilized in various hydro-climatological studies around the world and proven their applicability in different regions [9, 10]. The APHRODITE data (APHRO_MA_050deg_V1101) at 0.50°×0.50° grid scale from 1981 to 2007 was also utilized [11]. This gridded rainfall dataset has been prepared using gauged-based rainfalls and the angular distance weighting interpolation method was used for the gridding of rainfall observations [11].

3 Methodology

This study utilizes standard mathematical methods for filling the rainfall data gaps in SA-OBS after a careful review of literatures [6, 12]. Initially, the minor data gaps (<2%) in the SA-OBS data is filled by using the average of the nearest neighbouring rainfall grids. In terms of greater gaps (>2%), we applied four different methods to correct: (i) distance power (DP) method [12], (ii) distance power with high correlation coefficient (DPHCC) [12], (iii) linear regression (LR) method [12] and (iv) multiple linear regression (MLR) method [13]. Fig. 1 shows four grids that have data gaps greater than 2%.

The applicability of these methods in filling rainfall gaps is evaluated at one grid station (as testing grid with no data gaps) as shown in Fig. 1. We create some artificial data gaps for applying these methods and the relevant results are analysed and compared with original record. The data-gaps are created as per the reference of grids which have large data gaps (Green colour grids in Figure 1). Ten indicators are used for performance evaluation, such as coefficient of determination (R^2), percentage (%) of change, mean absolute error (MAE), root mean squared error (RMSE) and Akaike information criterion (AIC) [13]. For the spatial interpolation, we fix the cell resolution according to the actual grid resolution (i.e. 0.5°×0.5° grid scale), which is almost equivalent to 50 km^2. The standard rainfall extreme indices (REIs) such as Wet Days and Dry Days were generated as per the guidelines provided in previous studies [7]. The main purpose of generating REIs is to explore the rainfall extreme level changes in Malay Peninsula region.

4 Results and discussion

Table 1 shows the results of four different infilling methods (i.e. DP, DPHCC, LR, and MLR) for filling up rainfall data gaps present in the SA-OBS. Each method shows a reliable computation of missing rainfalls. Different statistical functions (as shown in Table 1) are utilized to test their applicability and strength over Malay Peninsula region, as the rainfall over this region is found highly inconsistent throughout the year.

Table 1. The statistical evaluation of infilling rainfall gaps based on four methods.

Statistical methods	SA-OBS (Original)	SA-OBS (Gaps)	DP	DPHCC	LR	MLR
Mean	9.70	9.57	9.55	9.62	7.97	9.80
% Change	NA	NA	-1.55	-0.81	-17.81	1.08
One Day Max	139.00	139.00	142.80	146.80	109.70	145.50
Std. Dev.	11.04	11.18	10.62	10.91	8.39	10.86
Std. Error	0.12	NA	0.12	0.12	0.09	0.12
R^2	NA	NA	0.97	0.98	0.86	0.98
MAE	NA	NA	1.08	1.00	2.45	0.92
MSE	NA	NA	3.09	2.73	16.82	2.29
RMSE	NA	NA	1.76	1.65	4.10	1.51
AIC	NA	NA	9048.20	8053.11	22654.89	6649.07

*NA = Not Applicable

Among all four methods, the DPHCC and the MLR methods give more satisfactory performance, as they recorded lower MSE, RMSE and higher R^2 than others. The DPHCC and MLR computed the lowest AIC values than other methods (Table 1). These two methods are also able to well capture the extreme rainfall events during the time series (1981-2007). Results showed that the methods which are selected for the rainfall gap filling could reasonably capture the one day maximum rainfall (1981-2007) as compared to SA-OBS (original). Therefore, these two methods are utilized for filling the rainfall data gaps at four grids of the SA-OBS data. Fig. 2 shows the results of the comparison of all datasets used in the study. Fig. 2(a) shows the difference in the daily mean rainfall (1981-2007) by taking the average of 27-year daily rainfall data. The comparison is conducted based on the same-time series length (i.e. 1981-2007) of all datasets across all grids. In Figs. 2(a), 2(b) and 2(c), the y-axes show the variations in daily mean, standard deviation and maximum rainfalls, respectively, across all grids. The PRINCETON rainfall shows relatively higher values than APHRODITE and SA-OBS in terms of mean, standard deviation, and maximum. The APHRODITE shows the lowest values for all on the contrarily. However, the PRINCETON data shows large variability in their ranges than other two datasets. Overall, in terms of average of all indicators, the SA-OBS is found relatively closer to PRINCETON.

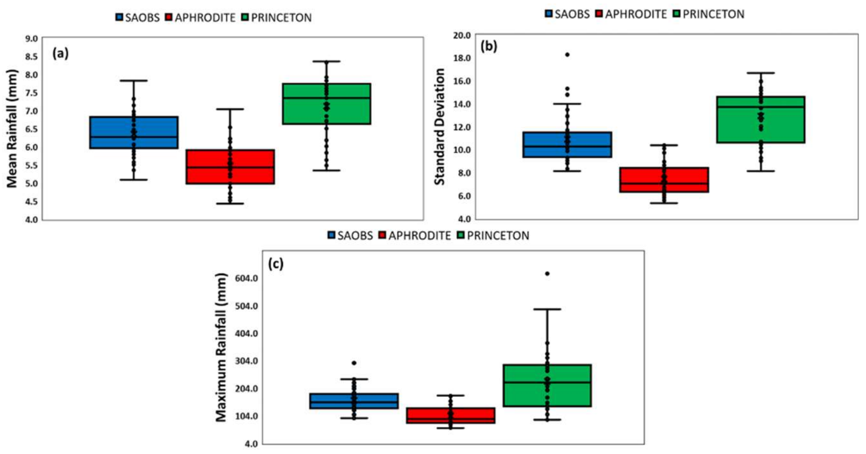

Fig. 2. The comparative statistical evaluation of SA-OBS, APHRODITE and PRINCETON rainfall datasets: (a) Mean Rainfall (daily mean), (b) Standard Deviation and (c) Maximum Rainfall.

The applicability and uncertainty with these datasets are also evaluated by using different statistical parameters as shown in Fig. 3. The results intend to explore the data characteristics of rainfalls based on their frequencies and mean behavior during the study period (1981-2007). Fig. 3(a) shows the spatial distribution of mean rainfall across the Malay Peninsula region. Overall, all plots indicate a notable variability across the whole region; especially, the APHRODITE shows slightly higher mean in the Northeastern side of Malay Peninsula. The SA-OBS and PRINCETON rainfalls are found relatively closer to each other. From Fig. 3(b), the DRY Days are recorded the highest in the case of PRINCETON and lowest in the one of APHRODITE as compared to SA-OBS. The Wet Days, as shown in Fig. 3(c), are recorded the highest in case of APHRODDITE, while SA-OBS and PRINCETON demonstrate a better consistency (Fig. 3c).

Fig. 4 shows the monsoon characteristics of Malay Peninsula region as explored by the three sources of datasets. The spatial plots are prepared based on the daily mean of rainfalls (1981-2007). The Malay Peninsula region is influenced by two monsoon systems including SASM and BWS, but it receives relatively more rainfalls during BWS. Fig. 4(a) clearly shows a higher mean of rainfalls across all the region in all datasets. The Malaya Peninsula region receives the lowest rainfall during JJA months as can be seen in Fig. 4(b). Overall, each dataset shows variability in their means across the entire region. PRINCETON and SA-OBS are found more comparable in SON and DJF, but APHRODITE and SA-OBS seems to be closer in MAM and JJA.

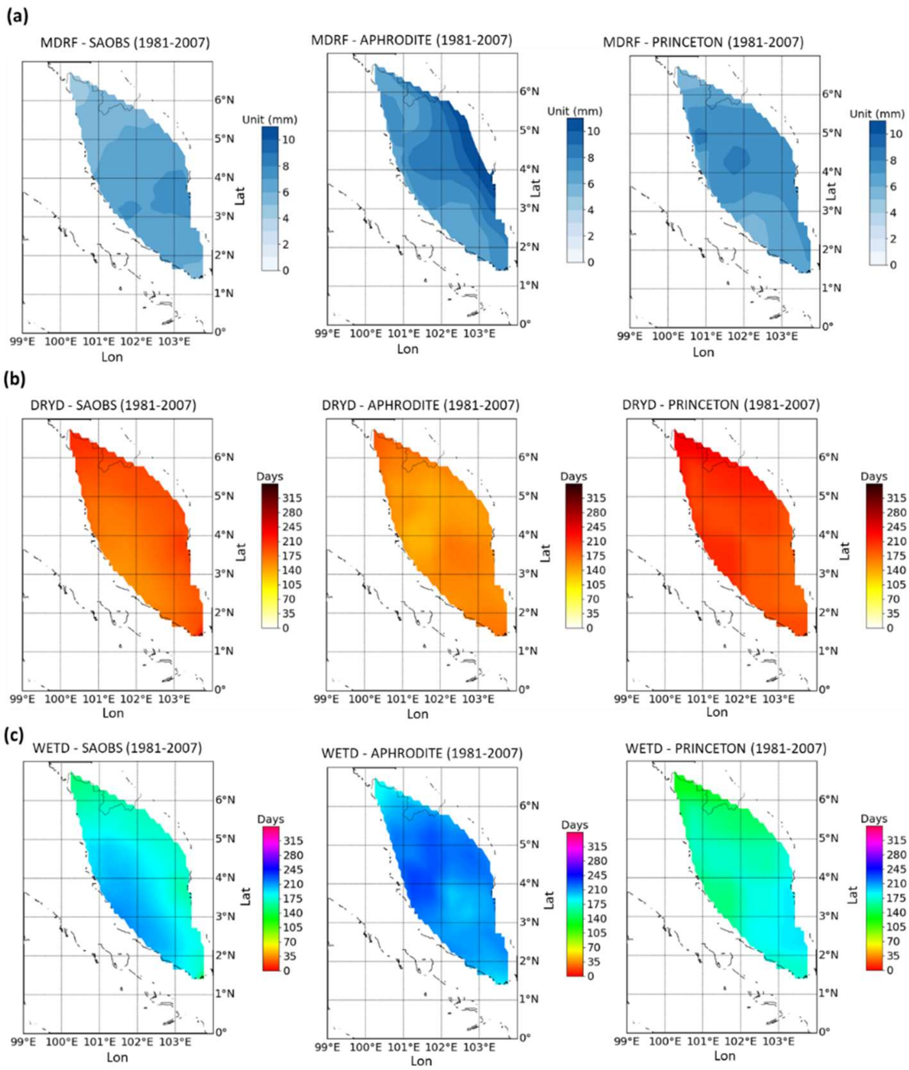

Fig. 3. The comparison of three datasets based on the (a) mean, (b) Dry Days (DRYD) and (c) Wet Days (WETD).

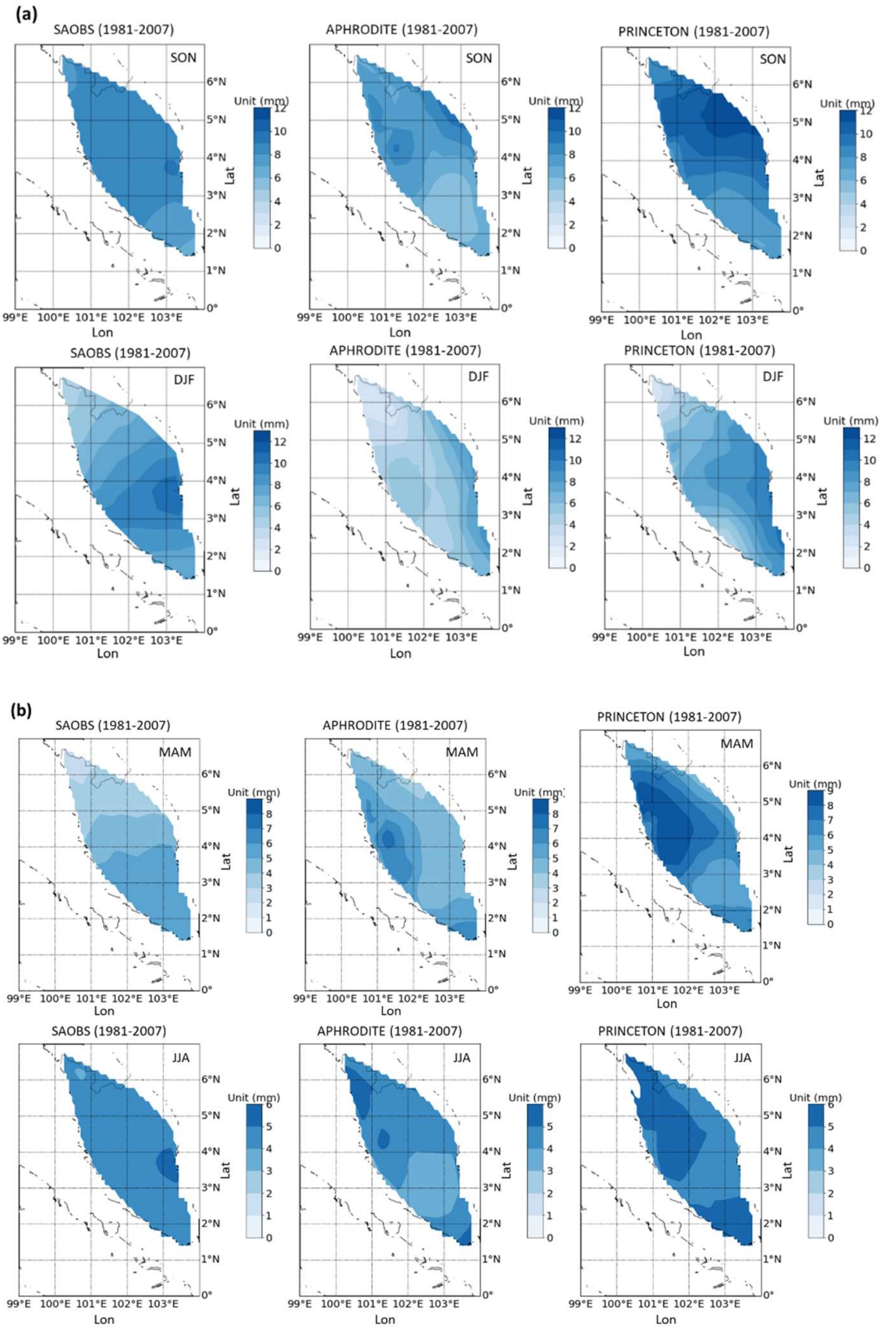

Fig. 4. The monsoon characteristics of Malay Peninsula region by three datasets in four seasons: (a) September-October-November (SON) & December-January-February (DJF) and (b) March-April-May (MAM) and June-July-August (JJA).

5 Conclusion

This study firstly carried out correction of the SA-OBS gridded rainfalls by filling the missing record in Malay Peninsula region. The availability of other standard global rainfall datasets, namely APHRODITE and PRINCETON, was also evaluated due to their wide applications around the world and in SEA. In this study, we utilized around 27 years daily time series rainfall datasets as per their availability and the data found sufficient to highlight our objectives, especially in SEA which has a large data limitations. The DPHCC and MLR are performed well in filling large rainfall gaps and similar approach may be applicable to other regions of the world in case of rainfall data scarcity. The seasonality of Malay Peninsula monsoon system has been explored and the spatial distribution of rainfall was analyzed over the region. The general findings were quite consistent from all datasets utilized in the study, although some variations in rainfall were found.

Acknowledgement

This project was supported by Start-Up Grant (M4081327.030) from School of Civil and Environmental Engineering, Nanyang Technological University, Singapore. We acknowledge the SA-OBS dataset and the data providers in the SACA&D project (http://saca-bmkg.knmi.nl). We are also thankful for the providers of PRINCETON (http://hydrology.princeton.edu/data.pgf.php) rainfall product and APHRODITE (http://www.chikyu.ac.jp/precip/english/products.html) rainfalls dataset.

References

1. V. Misra, S. DiNapoli, The variability of the Southeast Asian summer monsoon, Int. J. Climatol. 34(3): 893-901 (2014)
2. P. Xavier, R. Rahmat, W.K. Cheong, E. Wallace, Influence of Madden-Julian Oscillation on Southeast Asia rainfall extremes: observations and predictability, Geophys. Res. Lett. 41(12): 4406-4412 (2014)
3. E. Tyrlis, J. Lelieveld, B. Steil, The summer circulation over the eastern Mediterranean and the Middle East: influence of the South Asian monsoon, Clim. Dyn. 40(5-6): 1103-1123 (2013)
4. T.A. Woldesenbet, N.A. Elagib, L. Ribbe, J. Heinrich, Gap filling and homogenization of climatological datasets in the headwater region of the Upper Blue Nile Basin, Ethiopia, Int. J. Clim. 37(4): 2122-2140 (2017)
5. M.H. Kashani, Y. Dinpashoh, Evaluation of efficiency of different estimation methods for missing climatological data, Stoch. Environ. Res. Risk Assess. 26(1): 59-71 (2012)
6. R.S. Teegavarapu, A. Aly, C.S. Pathak, J. Ahlquist, H. Fuelberg, J. Hood, Infilling missing precipitation records using variants of spatial interpolation and data-driven methods: use of optimal weighting parameters and nearest neighbour-based corrections, Int. J. Clim. 38(2): 776-793 (2018)
7. P.V. Mandapaka, X. Qin, E.Y. Lo, Analysis of spatial patterns of daily precipitation and wet spell extremes in Southeast Asia, Int. J. Clim. 37(S1): 1161-1179 (2017)

8. E.J. Besselaar Van den, G. Van der Schrier, R.C. Cornes, A.S. Iqbal, A.M. Klein Tank, SA-OBS: a daily gridded surface temperature and precipitation dataset for Southeast Asia, J. Clim. 30(14): 5151-5165 (2017)
9. J. Sheffield, G. Goteti, E.F. Wood, Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, J. Clim. 19(13): 3088-3111 (2006)
10. A.M. El Kenawy, M.F. McCabe, A multi-decadal assessment of the performance of gauge-and model-based rainfall products over Saudi Arabia: climatology, anomalies and trends, Int. J. Clim. 36(2): 656-674 (2016)
11. A. Yatagai, K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, A. Kitoh, APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges, Bull. Amer. Meteorol. Soc. 93(9): 401-1415 (2012)
12. E. Sánchez, A. Avilés, E. Samaniego, Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes, Maskana, 5(1): 99-115 (2014)
13. M.T. Sattari, A. Rezazadeh-Joudi, A. Kusiak, Assessment of different methods for estimation of missing data in precipitation studies, Hydrol. Res. 48(4): 1032-1044 (2017)