Evaluation gridded precipitation datasets in Indonesia

T Wati1,2, T W Hadi1, A Sopaheluwakan2, and L M Hutasoit1

1Earth Sciences Graduate program, Faculty of Earth Sciences and Technology, Institute Technology of Bandung, Bandung, 40132, Indonesia

2Indonesia Agency for Meteorology Climatology and Geophysics BMKG, Kemayoran, Central Jakarta, 10610, Indonesia

Email: trinah.wati@gmail.com

Abstract. This preliminary study evaluates ten gridded precipitation datasets in Indonesia, namely APHRODITE, CMORPH, CHIRPS, GFD, SA-OBS, TMPA 3B42 v7, PERSIAN-CDR at 0.25°, moreover GSMaP NRT V06, GPM-IMERG (Early-Run) V06, and MSWEP V2 at 0.1° in the period of 2003 to 2015. The evaluation focuses on time series bias using metrics such as Mean Error, Coefficient of Variation, Relative Change (Variability), and Kolmogorov-Smirnov test (KS-test) at daily, monthly, seasonal, and annual time scales. The statistical relationship between the precipitation datasets with reference observational data use Taylor diagrams for evaluating the relative skill of the precipitation dataset. The study aims to evaluate the uncertainty of the precipitation datasets compared to rain gauge datasets. Time series bias of SA-OBS and MSWEP have the nearest value to zero as the best score. The relative skill of monthly rainfall based on rainfall typical shows that MSWEP outperformed in regions A and B, GPM-IMERG in C region. GPM-IMERG's relative skill is outperformed than other datasets at annual time scale in Region A and B, while TMPA 3B42 in Region C. The application of existing precipitation datasets is essential to cope with the limitation of rain gauge observations. This study implicates the development of precipitation products in the Indonesia region.

1. Introduction

Precipitation is a major and crucial component of the water cycle in the climate of the Earth [1]. Precipitation observation and recording that reliable and accurate are essential for the study of variability and climate change [2]. Rain gauge observations usually use for manual measurement of precipitation directly on the surface at a point scale. In Indonesia, the number of meteorological stations is still limited and sparsely distributed, especially in mountainous and inaccessible or remote areas that are likely not available. The application of existing gridded precipitation datasets can cope with the lack of rain gauge observations.

There are three categories of gridded precipitation datasets [3]: rain gauge based, satellite-based, and gauge-satellite combination datasets. The others are reanalysis datasets that use the assimilation scheme with climate models involving all observational data (radiosonde, satellites, buoy, planes, and ships). The combination of observation, satellite, and reanalysis complement the precipitation dataset variations. There have been numerous studies about gridded precipitation datasets that vary with time scales from hourly to decades [3], purpose, area coverage, spatial and temporal resolution [4].

This paper is a preliminary study evaluating existing global/regional precipitation gridded products over Indonesia. The objective is to evaluate the uncertainty of the precipitation datasets compared to rain
gauges datasets. The previous study of global comparisons found high uncertainties in the quantity and variability of precipitation at daily, monthly, seasonal, and annual time scales [2]. There are no best uniquely precipitation datasets in forcing hydrological models for all basins and depend on the basin characteristics [5]. Studies about the evaluation of gridded precipitation datasets in Indonesia have been carried out with several products [6] [7] [8]. This study focuses on the statistical relationship and time series bias between the precipitation datasets and observational data. For more specific, the analysis performs based on the classification of the rainfall region types in Indonesia [9].

2. Data and Methods

The study uses ten gridded precipitation datasets as presented in Table 1. The observation rain gauge datasets employ 133 rain gauges in Indonesia from 2003 to 2015 (Figure 1). The rainfall data are the same observed daily precipitation dataset used in Supari et al [22] up to 2012, with additional stations, periods and the same quality control analysis. All the observation rainfall data used in this study are homogenous. We use point to grid comparison approach. The grid location is the same as the point of the station site using the nearest neighbor method.

Table 1. Ten gridded precipitation products evaluated in this study

| No | Datasets                                                                 | Periods                  | Spatial Resolution |
|----|--------------------------------------------------------------------------|--------------------------|--------------------|
| 1  | APHRODITE (Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources) [10] | 1951-2007 (V1101)        | 0.25°              |
|    |                                                                          | 1998-2015 (V1901)        |                    |
| 2  | CHIRPS (Climate Hazards Group Infrared Precipitation with Stations) [11]  | 1981-present             | 0.25°              |
| 3  | CMORPH (Climate Prediction Center Morphing Technique) [12]                | 2002-2016                | 0.25°              |
| 4  | GFDv3 (Princeton Global Meteorological Forcing Dataset) [13]             | 1948-2016                | 0.25°              |
| 5  | PERSIANN-CDR (Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks-Climate Data Record) [14] | 1983-2015                | 0.25°              |
| 6  | SA-OBS (Southeast Asia Observation) [6]                                   | 1981-2017                | 0.25°              |
| 7  | TMPA 3B42v7 (Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis) [15] [16] | 1998-2019                | 0.25°              |
| 8  | GSMaP_NRT V06 (Global Satellite Mapping of Precipitation Near Real-Time) [17] [18] | 2000-present             | 0.1°               |
| 9  | GPM-IMERG V06 (Early Run) (Global Precipitation Measurement-Integrated Multi-satellite Retrievals for GPM) [19] [20] | 2000-present             | 0.1°               |
| 10 | MSWEPv2 (Multi-Source Weighted-Ensemble Precipitation) [21]              | 1979-2017                | 0.1°               |

The evaluation focuses on time series bias metrics Mean Error (ME) [22], Coefficient of Variation (CV), Root Mean Square Error (RMSE), Relative Change/RC (Variability), and Kolmogorov-Smirnov test (KS-test) daily, monthly, seasonal, and annual time scales [23] [24]. The metrics are a direct comparison of datum pairs on spatial/temporal location [25]. Table 2 describes formulas of all metrics with the best values are close to zero. The statistical relationship between the precipitation data sets (f) with reference observational data (r) use Taylor diagrams for evaluating the relative skill of the precipitation datasets. The Taylor Diagram [26] centralizes the error pattern information and assess different models comparison. Each point of the two-dimensional space in the Taylor diagram quantifies the degree of
correspondence of three statistics: the concentration of the difference in error/RMSE ($E^2$), the Pearson correlation ($R$), and standard deviation ($\sigma$). The Taylor diagram formula is as follows:

$$E^2 = \sigma_f^2 + \sigma_r^2 - 2 \sigma_f \sigma_r R$$  \hspace{1cm} (1)

Relative skills of monthly and annual precipitation datasets use Taylor Diagram based on rainfall typical regions in Indonesia [9].

Figure 1. The location of rain gauge stations with typical rainfall regions [9]

| Table 2. Formulas of the metrics |
|----------------------------------|
| **No.** | **Metric** | **Formula** | **Description** |
|-------|------------|-------------|-----------------|
| 1     | Mean Error (ME) | $ME = \frac{1}{N} \sum_{n=1}^{N} (y_n - x_n)$ | $y_n$ and $x_n$: time series of observed and precipitation dataset with N number of data [22]. |
| 2     | Coefficient of variation (CV) | $CV = \frac{|CV_m - CV_o|}{CV_o}$ | CV is a normalized measure of dispersion. $CV_o$ and $CV_m$ are coefficient of variation for the observation and a precipitation dataset [23] [24]. |
| 3     | Root mean square error (RMSE) | $RMSE = \frac{\left[\sum_{t=1}^{n}(p_t - f_t)^2\right]^{1/2}}{n}$ | $p_t$: time series data of a precipitation dataset, $f_t$: a time series data of the observation, and n: the number of time points. |
| 4     | Relative Change (RC) | $RC^t = \frac{P_{i+1} - P_i}{P_i}$ \hspace{1cm} $RC^t Bias = \frac{RC_{m} - RC_o}{RC_o}$ | Relative change (RC) evaluates relative difference or variability of precipitation dataset [23]. This metric reflects interannual variation of precipitation. |
| 5     | Kolmogorov-Smirnov Test (KST) | $D_{KS} = \max|F_o(x) - F_m(x)|$ | $F_o(x)$ and $F_m(x)$ are ECDF (empirical cumulative distribution function) of the observed data and a precipitation dataset (daily). $D_{KS}$ is maximum absolute difference between ECDF of two different datasets [23] [24]. |
3. Result and Discussion

Figure 2 shows the metric of ME all precipitation datasets, mean of daily ME range from -1.77 to 0.74 mm, while monthly ME from -61.9 to 24.6 mm. Precipitation datasets PERSIANN-CDR, CHIRPS, GFD, GPM-IMERG, and TMPA 3B42 (except MAM and SON) have positives ME indicating overestimates to observed data. On the contrary, APHRODITE, CMORPH, GSMaP_NRT, MSWEP, and SA-OBS are underestimated compare to observed data with negatives ME. The ranges of annual mean ME are from -742.8 to 295 mm. According to seasonal ME, DJF is the highest and the lowest on JJA. MSWEP has ME values that nearest to zero almost at all time scales except on DJF is GFD. The bias of CV represents the dispersion difference of a precipitation dataset from the observation. CV metrics at all timescales presented in Figure 3., which shows that SA-OBS has the lowest values that close to zero, and the highest is APHRODITE except at daily CV.

Daily RMSE metrics in Figure 4. range from 9.7 to 17.6 mm, while at monthly from 68 to 145 mm. Annual RMSE ranges from 433 to 1097 mm, and similar to ME, the range of DJF season is the highest while the lowest on JJA. The lowest RMSE is SA-OBS at all time scales and the highest is APHRODITE except on daily and SON. Figure 5. shows RC metrics distribution of observation and precipitation datasets over Indonesia. RC represents the interannual variation of precipitation, APHRODITE, PERSIAN-CDR, CMORPH, and GFD are underestimated in Borneo and overestimating in Java and Nusa Tenggara islands for APHRODITE and GSMaP_NRT. Precipitation dataset respectively has mean bias of RC metrics of 0.12 (APHRODITE), -0.05 (PERSIAN-CDR), -0.04 (CHIRPS), 0.03 (CMORPH), -0.05 (GFD), -0.04 (GPM-IMERG), 0.06 (GSMaP_NRT), 0.03 (MSWEP), 0.00 (SA-OBS) and -0.05 (TMPA 3B42) at all stations, with SA-OBS has mean RC value nearest to zero.

Figure 6. presents differences of daily ECDF (DKS) between precipitation datasets and observation over Indonesia, which have lower values (< 0.5) in Borneo and some parts of Sumatera, West Java, Central Sulawesi, and Papua. The lower value of DKS means the ECDF of the precipitation dataset is close to the observation data. Precipitation datasets mostly have a similar distribution of DKS, while PERSIAN-CDR and SA-OBS have more values > 0.75 than other datasets. DKS averages over Indonesia respectively are 0.58 (APHRODITE and CHIRPS), 0.67 (PERSIAN-CDR), 0.69 (SA-OBS), and 0.56 for others.

According to time series bias, uncertainty and variability of ten precipitation datasets at daily time scale range from 1 to 2 mm, while at monthly 25 – 61 mm and 295 up to 743 mm at annual time scale. DJF has the higher followed by MAM, SON, and JJA for seasonal time scales. The previous study [2] found that uncertainties among precipitation datasets up to 300 mm/year. SA-OBS and MSWEP have average metrics that nearest to zero. Since this analysis is a point-to-grid comparison of 133 location sites, not all datasets can capture the locations. The grids of CMORPH, GFD, GPM-IMERG, GSMaP_NRT, and TMPA 3B42 have 133 to compare, though there are some missing values in other datasets. SA-OBS and PERSIAN-CDR have 91 and 97 grids, and APHRODITE, CHIRPS, and MSWEP have fewer missing values (129, 128, and 131 grids). The evaluation of the precipitation dataset should consider the representativeness of the precipitation dataset grids.

The rainfall typical in Indonesia subdivided into three distinct climate sub-regions [9]: region A has monsoonal climate characteristics with a marked seasonal cycle near and south of the equator from south Sumatera to Timor Island, parts of Kalimantan, parts of Sulawesi, and parts of Papua. They are the rainy season centered on December to February and the dry season peaked from July to August [27] [28]. Region B covers the northern part of Sumatra and the Northwestern part of Kalimantan that have a less pronounced seasonal cycle where a high amount of rainfall almost throughout the year except for a decrease from June to August. Region C covers Maluku, and parts of Papua and part of Sulawesi (close to the western Pacific region) have an anti-monsoonal type, the rainy season from June to July and the dry season from November to February.
Figure 2. Metrics Mean Error (ME) at daily, monthly, seasonal and annual time scales

Figure 3. Metrics Covarian (CV) at daily, monthly, seasonal and annual time scales

Figure 4. Metrics RMSE at daily, monthly, seasonal and annual time scales
Figure 5. Distribution of RC Metric observation (a), APHRODITE (b), PERSIANN-CDR (c), CHIRPS (d), CMORPH (e), GFD (f), GPM-IMERG (g), GSMaP-NRT (h), MSWEP (i), SA-OBS (j) and TMPA 3B42 (k).
Figure 6. Absolute differences between precipitation dataset and observation of APHRODITE (a), PERSIANN-CDR (b), CHIRPS (c), CMORPH (d), GFD (e), GPM-IMERG (f), GSMaP_NRT (g), MSWEP (h), SA-OBS (i) and TMPA 3B42 (j).
Figure 7. Taylor Diagram of monthly mean precipitation based on rainfall region A (a), B (b) and C (c)

The statistical relationship uses decomposition schemes of the Taylor Diagram for three types of rainfall regions in Indonesia to further evaluate the relative skills of datasets and their performance on different types of rainfall. The position of each point appearing on the plot quantifies how closely that precipitation datasets pattern matches observations. The centred root-mean-square (RMS) difference between the gridded precipitation datasets and observed patterns is proportional to the distance to the point on the x-axis identified as "observed"[26]. Figure 7. shows the monthly mean precipitation of MSWEP has the nearest point to observation in Region A (Fig.7.a) and B (Fig.7.b), while Region C (Fig.7.c) is GPM-IMERG. The correlation coefficients of all datasets in Region A are more than 0.95 and in Region B are more than 0.90. Otherwise, the range of correlation coefficients in Region C is higher than A and B, from 0.70 to 0.99. The centered RMSE of MSWEP in Region A is about 97 mm/month, almost the same as observation, and in Region B is about 50 mm/month that lower than the observed. The standard deviation of the simulated pattern is proportional to the radial distance from the origin. The centered RMSE of GPM-IMERG in Region C is about 36 mm/month that higher than observed.

Meanwhile, in Figure 8. for annual precipitation, GPM-IMERG has the nearest point to observation in Region A and B (Fig.8.a and b, while TMPA 3B42 in Region C. The correlation coefficients in A range from 0.70 to 0.99, while in Region B and C are from 0.65 to 0.96. The centered RMSE of GPM-IMERG in Region A is about 310 mm/year that higher than observed, while in Region B is 190 mm/year lower than observed. TMPA 3B42 in Region C has centered RMSE about 410 mm/year lower than observation. Investigation uncertainty of precipitation datasets on different rainfall regions is essential for its application in a broad territory such as Indonesia.
Figure 8. Taylor Diagram of annual mean precipitation based on rainfall region A (a), B (b) and C (c)

4. Summary
Evaluation of existing gridded precipitation datasets is essential to evaluate uncertainties of the products. The uncertainty and variability of precipitation datasets comparing to observed data are crucial measures before its implementation. According to time series bias, uncertainty and variability of ten precipitation datasets at daily time scale range from 1 to 2 mm, while at monthly 25 – 61 mm and 295 up to 743 mm at annual time scale. DJF has the higher followed by MAM, SON, and JJA for seasonal time scales. SA-OBS and MSWEP have average metrics that nearest to zero. The relative skill of MSWEP monthly precipitation is outperformed in Region A and B, while GPM-IMERG in Region C. Meanwhile, for the annual time scale, GPM-IMERG outperform in Region A and B, TMPA 3B42 in Region C. The evaluation of gridded precipitation datasets based on typical rainfall region implicates the use of more than one precipitation dataset for climate and meteorology study in Indonesia.

5. References
[1] Eltahir E A and Bras R L 1996 Rev. Geophys. 34(3) 367
[2] Sun Q, Miao C, Duan Q, Ashouri H, Sorooshian S and Hsu K L 2018 Rev. Geophys. 56(1) 79
[3] Arakawa O and Kitoh, A 2011 Glob. Envi. Res. 15 109
[4] Beck H E, Vergopolan N, Pan M, Levizzani V, Van Dijk A I, Weedon G P, Brocca L, Pappenberger F, Huffman G J and Wood E F 2017 Hydrol. Earth Sys. Sci. 21(12) 6201
6. Acknowledgments
The authors are grateful to the Indonesia Endowment Fund for Education (LPDP) for funding this research. T Wati contributed as the main contributor to this paper. This paper is part of a dissertation with an improvement on characteristics of rainfall typical region in Indonesia.