Validation of Satellite and Ground Station-based High-Resolution Rainfall Gridded Datasets over the Philippines

J.C. ALBERT C. PERALTA *, GEMMA TERESA T. NARISMA

Atmospheric Science Program, Department of Physics,

School of Science and Engineering, Ateneo de Manila University, Philippines;

Regional Climate Systems Laboratory, Manila Observatory, Philippines

FAYE ABIGAIL T. CRUZ

Regional Climate Systems Laboratory, Manila Observatory, Philippines

(Submitted to *Journal of Hydrometeorology* November 2019)

* *Corresponding author address*: J.C. Albert C. Peralta I, Manila Observatory, Ateneo de Manila Campus, Loyola Heights, Quezon City, Philippines

Email: jc.peralta@obf.ateneo.edu

Current affiliation: Regional Climate Systems Laboratory, Manila Observatory

ABSTRACT

Characterizing the rainfall variability over the Philippines remains to be challenging due to the archipelago's complex shorelines and high terrain gradients, but only a few studies have been published partly due to the sparse spatial coverage of weather stations in the country. High-resolution gridded rainfall products can supplement the need for observed rainfall data in the mountainous regions and remote islands, but require careful validation with existing observations. This study compares four high-resolution, gridded datasets—APHRODITEv1101, CHIRPSv2, TRMM_3B42v7 and PERSIANN_CDR—with respect to 49 synoptic weather stations over the Philippines from 1998-2005. The performance of these datasets were assessed in terms of bias, distribution, and different statistical error metrics and skill scores across timescales and climate types. Results show that all the datasets were able to capture the basic climatology and to varying extents, spatial patterns of Philippine rainfall. TRMM_3B42v7 has the least overall average monthly bias and most closely resembles the rainfall distribution observed at weather stations, especially dry days and torrential rain days for the whole Philippines. APHRODITEv1101 performs best in terms of error metrics and skill scores but displays consistent underestimates. CHIRPSv2, on the other hand, best captures the seasonal rainfall peaks in the different climate types in the Philippines but is prone to larger errors. Lastly, PERSIANN_CDR shows generally poor metrics and rainfall distributions, in comparison to the other datasets. These key findings are used to identify possible research applications in the Philippines that are best suited for each dataset.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

1. Introduction

Rainfall plays a major influence on climate variability in the Philippines. At seasonal scales, rainfall is driven by the large scale flow of the Asian summer monsoon (e.g.Murakami and Matsumoto 1994), and the Asian winter monsoon (e.g.Juneng and Tangang 2010). Tropical cyclones and low-pressure systems originating from the Pacific Ocean also bring extreme rainfall to the country (Cinco et al. 2016; Bagtasa 2017). Other well-studied global scale phenomena such as the El Niño–Southern Oscillation (ENSO) (Hilario et al. 2009) and the Madden Julian Oscillation (MJO) (Pullen et al. 2015) also modify the annual rainfall amount and distribution in the country.

There are four climate types in the Philippines, as defined by the Modified Coronas climate classification (Coronas 1920; Flores and Balagot 1969; Kintanar 1984). These climate types are determined based on rainfall variation observed from ground stations over the whole country, and are characterized as follows:

a) Type 1 – These areas are located over the western side of the country facing the West Philippine Sea, and have two pronounced seasons: wet (typically above 100 mm month^{-1}) associated with the Asian summer monsoon from May to September, and dry (typically below 100 mm month^{-1}) from November to April.

b) Type 2 – Situated over the eastern side of the country facing the Pacific Ocean, these areas characteristically do not have a dry season, but display a pronounced rainfall peak observed from October to January. This peak is associated with the Asian winter monsoon and Western Pacific tropical cyclones.

c) Type 3 – Located immediately east of Type 1, this type is either over inland areas or islands within the archipelagic seas. These areas have no pronounced

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

peaks and have less total annual rainfall than either Type 1 or Type 2, but they record similar but slightly wetter dry months compared to Type 1.

d) Type 4 – These areas lie west of Type 2 with annual rainfall comparable to Type 3 but more or less evenly distributed over all the months. However, the profile resembles that of Type 2 because of a similar absence of a pronounced dry season.

This classification is a good survey of Philippine rainfall variability as the majority of the features described was also present when an objective clustering method was applied to rainfall data from recent station and satellite-based observations (Corporal-Lodangco and Leslie 2017). It is also proven to be useful in characterizing long-term rainfall trends and extreme indices (Cruz et al. 2013; Villafuerte et al. 2014). The spatial extent of each climate type is presented in Figure 1.

Given the country's high vulnerability to the effects of climate change (Kreft et al. 2014), the pressing public need for timely and accurate rainfall data for monitoring and research must be continually addressed. High-resolution gridded rainfall data products derived from satellite imagery and interpolated rainfall station networks are becoming increasingly accessible in recent years. Because of their high spatial and temporal resolution and coverage, these gridded datasets show promise as reliable and cost-effective sources of rainfall information, particularly in inaccessible rural regions where in situ measurements are either sparse or absent. The broad spatial coverage of these gridded rainfall datasets is also critical in evaluating climate models, an essential first step to account for uncertainties relevant to end-users and to provide insight into future model developments. However, these gridded rainfall datasets need to be validated first with observed station data to determine their capability to capture different rainfall characteristics that are crucial for specific applications.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

Jamandre and Narisma (2013) earlier validated a selection of three gridded rainfall datasets in the Philippines, namely, the Asian Precipitation—Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) version 1101(Yatagai et al. 2012), Tropical Rainfall Measuring Mission (TRMM) 3B42 version 6 (Huffman et al. 2007), and the Climate Prediction Center morphing method (CMORPH) (Joyce et al. 2004). These datasets were compared with rainfall data from eight synoptic weather stations of the national weather agency, Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA), from 2003-2005. Their results showed that the correlation with station data of the ground-station interpolated APHRODITE is consistently better than the satellite-derived TRMM and CMORPH for all locations considered, especially the northern and eastern regions of the Philippines. They have also noted that APHRODITE and CMORPH recorded high errors for low rainfall and that TRMM gave more accurate estimates for higher rainfall values.

This paper extends the rainfall validation of Jamandre and Narisma (2013) by increasing the number of synoptic weather stations, lengthening the time period, as well as considering other high-resolution rainfall datasets. The goal is to help recommend potential applications that are best suited for each dataset. Section 2 of this paper describes the geographical features of the Philippines, followed by details about the datasets and methods used. The results of the different validation procedures are discussed in Section 3. Finally, Section 4 presents a summary of key findings and applications recommended per dataset.

2. Data and Methods

2.1 Study Domain

The Philippines is an archipelagic country located in the northeastern part of Southeast Asia (Figure 1). It is bounded by the Luzon Strait to the north, the Pacific Ocean to the east, the West Philippine Sea to the west and the Celebes Sea to the south. The country is composed

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

of over 7000 islands divided into three major groups: Luzon (composed of Luzon Island, Mindoro, Palawan and neighboring small islands), Visayas (composed of major islands around the center of the archipelago) and Mindanao (composed of Mindanao Island and neighboring small islands) (Central Intelligence Agency 2019). The topography within these islands is quite complex—narrow mountain ranges, lowlands, and jagged coastlines—which also influences local-scale rainfall.

2.2 Ground Station Data

Daily rainfall data from synoptic stations of PAGASA were used as the basis for validation (Villafuerte et al. 2014; Cinco et al. 2014). Given the varying temporal coverage across the different stations, the longest common time period of station data with all four gridded datasets was used in this study, which is from 1998 to 2005. After discarding stations with more than 20% missing or invalid daily values, only 49 out of 52 available stations were included in the analysis (Supplementary Table 1). These stations are well distributed across the archipelago, covering all climate types: Type 1 (16 stations), Type 2 (12 stations), Type 3 (12 stations) and Type 4 (9 stations) (Figure 1).

2.3 High-Resolution Gridded Rainfall Datasets

Four gridded rainfall products were selected for validation due to their high resolution and documented performance in other tropical regions (Usman et al. 2018; Dinku et al. 2018; Diem et al. 2014; Prasetia et al. 2013). These are APHRODITE version 1101 (APHRODITEv1101) (Yatagai et al. 2012), the Climate Hazards Group Infrared Precipitation with Stations version 2 (CHIRPSv2) (Funk et al. 2015a), the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN_CDR) (Ashouri et al. 2015) and the TRMM 3B42 version 7 (TRMM_3B42v7) (Huffman et al. 2007).

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

APHRODITEv1101 was produced by interpolating rainfall measured by a network of 5,000–12,000 stations from national hydrological and meteorological services institutions all over the Asian continent, in addition to rainfall data provided by the World Meteorological Organization (WMO) Global Telecommunication System (GTS) (Yatagai et al. 2012). The data is available for the domains of Monsoon Asia (60°E–150°E, 15°S–55°N), Middle East (15°E–65°E, 25°N–45°N), and Russia (15°E–165°W, 34°N–84°N), covering 57 years from 1950 to 2007. The interpolation algorithm involves a multistep interpolation process using the daily station observation weighted by its ratio to monthly and daily climatologies at a 0.05° resolution grid and finally regridding to 0.25° and 0.5° resolution products.

CHIRPSv2 is a rainfall gridded dataset covering 1981 to the present, available at 0.05° and 0.25° resolution based on infrared cold cloud duration (CDD) satellite observations and station-based data. The satellite-only product, CHIRP, takes in data from two infrared imagery sources: the Globally Gridded Satellite (GridSat) archive (1981-2008) produced by NOAA's National Climate Data Center and the 2000–present NOAA Climate Prediction Center dataset (CPC TIR). The rainfall estimates are produced by using a fixed CCD threshold and calibrated by 0.25° resolution TRMM 3B42 pentadal precipitation (2000-2013). This is blended with station-based monthly rainfall climatology of the CHPClim dataset (Funk et al. 2015b) using a modified inverse distance weighting algorithm described in (Funk et al. 2015a).

The PERSIANN_CDR dataset contains daily rainfall estimates at 0.25° resolution over 60°S-60°N from 1983 to the near-present (Ashouri et al. 2015). The precipitation estimation technique combines the PERSIANN algorithm on GridSat-B1 infrared satellite data and an artificial neural network estimation method trained using the National Center for Environmental Prediction (NCEP) high-resolution Doppler radar data. The resulting estimate is adjusted such that it matches the Global Precipitation Climatology Project (GPCP) monthly product version 2.2 (Adler et al. 2003) when regridded to the latter's 2.5° resolution.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

The TRMM_3B42v7 dataset is produced by combining data from the Precipitation Radar (PR) and TRMM Microwave Imager (TMI) instruments aboard the TRMM satellite in an algorithm described in Huffman et al. (2007). Like PERSIANN_CDR, this dataset is also calibrated by monthly gauge data from the Global Precipitation Climatological Center (GPCC) (Rudolf 1993) and the Climate Assessment and Monitoring System (CAMS) monthly rain gauge analysis (Xie et al. 1996). The TRMM mission came to an end on April 8, 2015 and was succeeded by the Global Precipitation Mission (Huffman 2017). TRMM_3B42v7 covers global data within 50°N-50°S from 1998 up to near-present, but it is recommended that this product version is used only for January 1998 to September 2014 to ensure homogeneity.

This study used the 0.25° resolution variant for all the four datasets mentioned. The PAGASA station rainfall dataset was compared with the gridded rainfall datasets at the grid containing or nearest the station location. For analysis involving the monthly and seasonal scales, a month aggregate is discarded for each dataset if there are at least 5 days of missing data within it—removing a total of 19 months distributed across 11 stations. It should be noted that all these datasets have incorporated PAGASA station data in their estimating algorithms at various levels—either as an explicitly listed input in the case of APHRODITEv1101, or as input to the rain-gauge-based datasets that were used in calibration or bias-correction, in the case of the other three satellite-derived gridded datasets.

2.4. Comparing climatological rainfall profiles and rainfall event category distributions

For each climate type, the median monthly and annual totals, as well as the 25th percentile (P25) and 75th percentile (P75), rainfall values of gridded datasets were first compared with PAGASA to initially identify well and poor performing temporal features of the gridded datasets per climate type. Next, the daily rainfall events were classified into four categories, namely: dry days, light to moderate rain days, heavy rain days, and very heavy rain days (Table 1). The thresholds of these categories were based on the definition of extreme

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

196 rainfall indices as recommended by the World Meteorological Organization (2018). The
197 monthly and multi-year frequency and contribution to the total rainfall of these categories as
198 represented in the gridded datasets were compared with those observed from PAGASA
199 stations.

200 *2.5. Analyzing the rainfall bias*

201 The accuracy of the gridded dataset rainfall is quantified by the daily bias *(B)*, defined
202 as the difference between the gridded dataset value *d* and the PAGASA value *o* for each day *i*
203 (Eqn. 1), and the daily bias fraction *(BF)*, defined as the daily bias divided by the PAGASA
204 value (Eqn. 2)

205 $$B_i = d_i - o_i \tag{1}$$

206 $$BF_i = \frac{d_i - o_i}{o_i} = \frac{B_i}{o_i} \tag{2}$$

207 where $BF = 0$, if the gridded dataset perfectly replicates the station value; $BF = 0.5$, if
208 the gridded dataset is half of station value; $BF = 1$ if double; $BF = 2$, if triple, and so on. A
209 quantile-quantile (Q-Q) plot is also drawn to show the deviations in the percentile value
210 distributions of the datasets with respect to PAGASA rainfall, providing an initial survey of
211 the expected bias across all rainfall values. Lastly, to show the bias of each dataset in more
212 detail, the average monthly bias is also shown with an accompanying split distribution of its
213 contributing daily wet and dry biases.

214 *2.6. Describing the spatial patterns of bias, correlation, and frequency of extremes*

215 Seasonal maps comparing each gridded dataset at each PAGASA station location were
216 drawn to determine the reliability of each gridded dataset over the whole Philippines. Aside
217 from showing the seasonal bias value, the bias maps emphasize locations where the biases are
218 consistently observed. The seasonal bias over a location is said to be consistent if a simple
219 majority (50%+1) of its constituent monthly biases have the same sign with the seasonal mean

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

bias. Maps of the daily correlation of the gridded datasets with respect to PAGASA were also done over the whole time period to identify areas where rainfall estimates better/poorly fit the observed. Lastly, maps of the frequencies of dry and very heavy rain events in PAGASA and the gridded datasets were compared to identify areas over which the better performing gridded datasets could potentially be used for extremes analyses.

2.7. Calculating error metrics and skill scores

The reliability of a gridded dataset's performance in increasingly longer timescales is important for applications that require the use of multiple-scale metrics like climate model evaluation and the development of extreme indices. For this analysis, the daily, monthly and annual rainfall totals of the datasets were compared with PAGASA station values using the root mean square error (RMSE), mean absolute error (MAE), linear regression coefficient (R^2), and the Kolmogorov Smirnov distance (KS), averaged over all stations. These are given by the following expressions:

$$RMSE = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(d_i - o_i)^2\right)} \tag{3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|d_i - o_i| \tag{4}$$

$$R^2 = \left(\sum_{i=1}^{n}\frac{(d_i - \bar{d})(o_i - \bar{o})}{(n-1)\sigma_d\sigma_o}\right)^2 \tag{5}$$

$$KS = \max_k(|D_k - O_k|) \tag{6}$$

where d_i is the gridded dataset value and o_i is the PAGASA station value, σ_d and σ_o are their respective standard deviations, D_k and O_k are their cumulative distribution function values at percentile value k, and n is the total number of timesteps.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

The RMSE (Eq. 3) is the standard deviation of the differences between the station and gridded dataset rainfall. Since this metric squares the difference, large deviations from station values have bigger contributions to the error. The MAE (Eq. 4) is similar but with all differences taken as absolute values and weighted equally in the score as a simple average. The R^2 (Eq. 5) represents the proportion of variance in the station values that is explained by the gridded dataset. Lastly, KS (Eq. 6) is a nonparametric measure quantifying the similarity of the rainfall distributions, given as the maximum distance between the cumulative distribution functions of the station and the gridded rainfall.

Furthermore, skill metrics based on a contingency table were used to assess the accuracy of the gridded datasets with respect to station values in the daily timescale. This analysis was performed using four event categories previously described in Table 1 as thresholds. Considering the historical impacts of extremely high rainfall events in the Philippines, the ability of gridded datasets to capture very heavy rainfall is of special interest in this study. Thus for this analysis, the category for very heavy rainfall days was further divided into three subcategories consistent with those used by Jamandre and Narisma (2013): (1) 20-50 mm, (2) 50-100 mm, and (3) greater than 100 mm.

Daily events that satisfy the category thresholds were aggregated to obtain the number of hits, false alarms, misses, and correct negatives for all years considered. From a roster of skill metrics as documented by WMO (Mason 2016) and regional studies (Zhang et al. 2000; Ghelli 2009), the following were selected: the Frequency Bias Score (FBS), Probability of Detection (POD), False Alarm Rate (FAR) and the Peirce Skill Score (PSS). The formulae are given below, where A, B, C, and D represent hits, false alarms, misses, and correct negatives, respectively.

$$FBS = \frac{A + B}{A + C} \tag{7}$$

$$POD = \frac{A}{A + C} \tag{8}$$

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

$$FAR = \frac{B}{B+D} \quad (9)$$

$$PSS = \frac{A}{A+C} - \frac{B}{B+D} \quad (10)$$

The FBS (Eq. 7) compares the event detection frequency in the station values with the gridded rainfall estimates, demonstrating underdetection (overdetection) if the value is less (greater) than 1. The POD (Eq. 8) is the proportion of observed rain events correctly detected by the gridded dataset, which should ideally be close to unity. The FAR (Eq. 9) compares the false alarms from all detected event instances, with event detection skill becoming indistinguishable from random chance as this value approaches unity. Lastly, the PSS (Eq. 10) is the difference between POD and FAR ranging from 0 (no skill) to 1 (perfect skill). Because it measures the overall ability of the dataset to correctly classify events and non-events, it has shown practical advantages in forecast evaluation compared to the more widely used Heidke Skill Score, which compares accuracy to random chance (Appleman 1960; Woodcock 1976).

3. Results

3.1 Climatological rainfall profiles

The spatial patterns of the average seasonal rainfall (1998-2005) from PAGASA stations (Figure 2) illustrate the seasonal rainfall description of the climate types of the Philippines. During December-February (DJF) (Figure 2a), the increased activity of the Asian winter monsoon produces a sharp west-east spatial contrast between the dry (0-50 mm) Type 1 and wet (800-1200 mm) Type 2 regions. In contrast, the driest conditions (0-500 mm) over the whole country follow during March-May (MAM) (Figure 2b) but with a small rainfall peak over the eastern and western coasts due to easterlies and onset of the summer monsoon, respectively. When the summer monsoon peaks during June-August (JJA), Type 1 regions over the western coast of the Philippines receive the highest rainfall (700-1000 mm). This is considerably lower than the peak in DJF over the eastern coast, but there is high rainfall over a much wider area (Figure 2c). This rainfall pattern persists during September-November

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

(SON) but with higher rainfall over the Type 2 regions as the winter monsoon and the Western Pacific typhoon activity increase (Figure 2d). Lastly, stations within Type 3 and Type 4 regions are observed to have relatively low rainfall variation (300-700 mm) throughout the year.

Although there are large deviations, the monthly profile of all the gridded datasets generally captured the shape of the station profile for every climate type (Figure 3). Most datasets also showed considerable difficulty in replicating the spread in the monthly values. For Type 1 (Figure 3a-d), CHIRPSv2 performs best with annual and monthly medians closest to station values, although the rainfall median from June-February is slightly higher than PAGASA median. Meanwhile, TRMM_3B42v7 overestimates the annual median by a similar magnitude as CHIRPSv2, but becomes larger during September-February. On the other hand, APHRODITEv1101 clearly underestimates Type 1 annual rainfall because of the lower summer monsoon peak from May-September. The least accurate profile for Type 1 is from PERSIANN_CDR as it has underestimated the peak on June-September and notably overestimates the median beyond the PAGASA monthly spread.

For Type 2 (Figure 3e-h), all datasets perform well during the relatively drier months of May-October, but CHIRPSv2 is the only dataset that captures the rainfall peak from September-February, and the spread for all months. The rest of the datasets consistently show difficulty in capturing this peak, which translates to the annual rainfall below the annual P25 of PAGASA. TRMM_3B42v7 shows a rainfall median peak close to PAGASA for October-December but largely underestimates the median in January-March. On the other hand, APHRODITEv1101 consistently underestimates the Type 2 profile peak for the entire winter monsoon season from October-February and March-April. Lastly, PERSIANN_CDR has the

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

poorest performance among the datasets with an annual median deficit because of a very low median during October-April that even falls below or lie close to the P25 of PAGASA.

For intermediate Type 3 (Figure 3i-l) and Type 4 (Figure 3m-p), APHRODITEv1101 performs the best among all datasets. Its profile and spread are the closest to PAGASA, albeit underestimated, similar to the other climate types. The other three datasets, in contrast, produce consistent median overestimates for these regions, which are mainly composed of inland valleys and interior islands.

3.2 Daily rainfall event category distribution

Among all datasets, TRMM_3B42v7 (Figure 4e,j) estimated the closest rainfall partitions to PAGASA (Figure 4a,f) in both frequency and contribution . The frequency of dry days is overestimated from November to April and underestimated for the remaining months, but never deviated beyond 5% of PAGASA. The rainfall contributions of all rainfall categories in TRMM_3B42v7 were also accurate to within 5% of PAGASA for all months, except very heavy rainfall days, as they are slightly overestimated from November to February by as much as 10%. CHIRPSv2 also captures the general trend of the frequency and contribution of the station rainfall (Figure 4c,h), but it exceeds the frequency of rainfall days of PAGASA from June to October by around 3-10%. In terms of contribution, the dataset best captures the station rainfall frequency and contribution from November to February among all datasets, which agrees with its good performance in capturing the peak rainfall profile of Type 2, as discussed in Section 3.1.

APHRODITEv1101 and PERSIANN_CDR did not represent the frequency and contribution of rain categories as well as TRMM_3B42v7 and CHIRPSv2. The consistent underestimates in APHRODITEv1101 for Type 1 and Type 2 regions could be explained by its evident overrepresentation of light-to-moderate rainfall days for the whole year, which could be as high as double the frequency in June-December and almost triple the observed

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

contribution for March-May in PAGASA (Figure 4b,g). Heavy rainfall days also exceed frequencies and contributions in PAGASA by about 40-50%. Among all the datasets, only APHRODITEv1101 was unable to represent the percentage of no-rain days in the dry days. PERSIANN_CDR also has a similar excess of light-to-moderate rainfall days but is slightly close to PAGASA (Figure 4d,i). However, PERSIANN_CDR has the highest contribution of heavy rainfall days and the lowest contribution of very heavy rainfall days during March-September for all the datasets, which could be the reason for its low rainfall profile peaks during this period.

3.3 Rainfall Bias

The excesses and deficits in the frequency and contribution of each daily event category also manifest in the daily rainfall bias and bias fraction (see Section 2.5), which reveal differences between the ground-station-interpolated dataset APHRODITEv1101 and the other three satellite-derived datasets. Large discrepancies in median and spread of bias fractions for dry days and light-to-moderate rainfall days, with respect to PAGASA, indicate that all datasets have difficulty in capturing low rainfall (Figure 5a). This is expected as rainfall rates less than 2 mm h^{-1} is still below sensitivity thresholds of typical radar-based rainfall measurements (Arulraj and Barros 2017). For dry days, the median bias fractions of APHRODITEv1101, CHIRPSv2, and PERSIANN_CDR at BF = 6.0, 1.5 and 8.0, respectively, show that the rainfall estimates of the datasets typically exceed twice the observed value. The exception in this category is TRMM_3B42v7, which has a negative median bias fraction (-1.0), indicating that it estimates dry days more often as a no-rain day. For light to moderate rainfall, the performance of all the datasets was seen to improve in terms of the median BF but still retains a wide spread. For heavy and very heavy rainfall days, all datasets underestimate observed rainfall, with reduced bias fraction spreads that is most probably due to fewer events in these categories. The satellite-derived datasets also have relatively larger BF spreads. On the other hand, the ground-

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

station-interpolated APHRODITEv1101 recorded the least *BF* and spread among all the datasets, but this could be due to its underestimation of the frequency of high rainfall events, as discussed in Section 3.2.

The rainfall Q-Q plot of the datasets (Figure 5b) shows that APHRODITEv1101 consistently underestimated rainfall across all percentile values by about 80-160% of PAGASA, which is consistent with its rainfall profile and distribution as discussed in Sections 3.1 and 3.2. On the other hand, the satellite-derived datasets tend to record higher than PAGASA percentile values up to a threshold specific to the dataset, then underestimate succeeding higher percentiles. TRMM_3B42v7 achieved the closest percentile distribution to PAGASA rainfall, with the widest range of rainfall percentiles (0.6-1.0 mm and 12.0-100.0 mm) within 10% of PAGASA. Meanwhile, CHIRPSv2 rainfall percentiles fell within 10% of PAGASA only for the percentiles within 30-50 mm, and PERSIANN CDR for a much lower range of 8.5-15 mm.

The bidirectional box plots of the rainfall bias (Figure 6) concisely reveal features of the bias distribution that will not be seen if only the average is presented (see Figure 7). APHRODITEv1101 has consistently recorded frequent (i.e., 69-77% of all days) positive daily biases for all months (Figure 6a). This is expected from the significantly larger (smaller) light to moderate rain days (dry day) partitions noted previously in Section 3.2. However, the monthly median magnitudes of positive biases are remarkably small (0.5-2.0 mm day^{-1}) compared to the median magnitudes of negative biases (2.6-6.0 mm day^{-1}), causing an overall average underestimation (5.0-10.0 mm day^{-1}) of the dataset for all months. This underestimation is consistent with Jamandre and Narisma (2013), and is also observed in regions, e.g. Pakistan (Ali et al. 2012). It can also be noted that APHRODITEv1101 has the smallest spread of biases among all datasets, which is most likely due to a conservative interpolation algorithm minimizing deviations from the rain gauge values (Yatagai et al. 2012).

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

With similar partition deficits to APHRODITEv1101 for days with recorded rainfall, PERSIANN_CDR shares the same consistent albeit less frequent positive daily biases for all months (54-71%), except in January (Figure 6c). The monthly median magnitudes for both positive and negative biases are almost twice as large as APHRODITEv1101 for all months. Furthermore, the overall mean negative biases are higher than APHRODITEv1101 with a wider spread. The very high dry bias (10.0-13.0 mm day^{-1}) during the summer monsoon months from June to September is most prominent for this dataset. This large dry bias is also present in other regions e.g., Colombia (Dinku et al. 2010) and eastern China (Liu et al. 2015b).

CHIRPSv2 (Figure 6b) and TRMM_3B42v7 (Figure 6d) are similar in bias distribution, but the former has a slightly larger median daily bias and a wider spread in daily bias for all months. Rainfall is generally underestimated during the summer monsoon months from May to October in both CHIRPSv2 (8.0-10.0 mm day^{-1}) and TRMM_3B42v7 (1.0-5.0 mm day^{-1}). During these months, positive daily biases also occur more often (53-63%), but the larger spread of less frequent negative biases dominates the overall mean monthly bias for both datasets. For the winter monsoon months of November to February, TRMM_3B42v7 has consistently overestimated rainfall (1.0-3.0 mm day^{-1}), which is similar to CHIRPSv2 in December and January (1.0-4.0 mm day^{-1}), but with a slight overestimation in March (1.5 mm day^{-1}).

Overall, TRMM_3B42v7 performed best in terms of bias because it has the lowest monthly mean biases. This low bias is also seen in other regions such as Brazil (dos Reis et al. 2017), Uganda (Diem et al. 2014) and eastern China (Liu et al. 2015). Interestingly, the positive and negative bias medians for all months are also almost of the same magnitude. Although positive biases occurred slightly more often from April to October, and negative biases for the rest of the year, the difference is at most 9%, which is relatively small compared to the other

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

datasets. This could be expected given the closeness of the event partitions and percentile distributions with PAGASA.

3.4 Spatial patterns of bias, correlation, and frequency of extremes

The signs of the biases of the datasets vary geographically across the Philippines (Figure 7). In December-February (Figure 7a-d), all datasets show mostly consistent dry biases over eastern Philippines where the winter monsoon rainfall peak is expected. On the other hand, the relatively drier western Luzon and central Mindanao show consistent wet biases at varying magnitudes, most evident for PERSIANN_CDR. During March-May, low biases over the whole Philippines were recorded (Figure 7e-h), but the monthly signs are mostly inconsistent. APHRODITEv1101 and CHIRPSv2 have the smallest bias magnitudes, while PERSIANN_CDR and TRMM_3B42v7 show some interspersed strong dry and wet biases over some islands. In June-August (Figure 7i-l), consistent large dry biases were found in the summer monsoon rainfall peak over western Philippines, most noticeable for APHRODITEv1101 (Figure 7i). The satellite-derived datasets (Figure 7j-l) also have dry biases over western Luzon but show wet biases over the interior islands of western Visayas and northern Mindanao. Finally, for September-November (Figure 7m-p), small biases were recorded in central and southern Philippines. However, bias magnitudes varied widely for Luzon: strong dry biases were seen in APHRODITEv1101 (Figure 7m), while the rest of the datasets show large biases with a mixture of signs (Figure 7n,o,p).

In terms of the spatial distribution of daily correlation values of the datasets with respect to PAGASA, APHRODITEv1101 (Figure 8a) recorded the highest values for majority of the locations considered. This could be expected because the dataset uses PAGASA station data in its interpolation method. The highest correlations occur over southwestern Luzon, central-eastern Visayas and northeastern Mindanao. On the other hand, the satellite-derived datasets show consistent lower correlations with observed values (Figures 8b-d). For these datasets,

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

correlation over southern Luzon is relatively higher compared to the rest of the country while Mindanao has relatively lower values, as was also previously seen in Jamandre and Narisma (2013). Interestingly, all satellite-derived datasets have higher correlation than APHRODITEv1101 over Batanes Islands at the northernmost part of the country, Palawan and southwestern Mindanao. For TRMM_3B42v7, more locations over western coasts of the country have higher correlation values than APHRODITEv1101, which could be due to the former's better representation of summer monsoon rainfall. Thus, these geographical locations could possibly show the areas where pure interpolation algorithms may not be adequate and where satellite input can help better estimate the observed rainfall.

The capability of these high-resolution datasets to capture rainfall extremes adequately for impact studies, such as drought and torrential rains due to tropical cyclones and monsoon events, is also examined. Both TRMM_3B42v7 (Figure 9e) and CHIRPSv2 (Figure 9c) performed well in obtaining the annual spatial distribution of observed PAGASA dry days; thus, these datasets are well suited for use in drought analysis. On the other hand, APHRODITEv1101 (Figure 9b) and PERSIANN_CDR (Figure 9e) lacked almost half of the observed dry days for the whole country. For the extreme rainfall categories of 20-50 mm, 50-100 mm, and >100 mm, TRMM_3B42v7 (Figure 9j,o,t) captured the spatial patterns quite well compared to the other datasets. CHIRPSv2 (Figure 9h,m,r) slightly overcounted the 20-50 mm and 50-100 mm categories over eastern Philippines but was unable to capture the spatial extent of the eastern peak for >100 mm. In contrast, APHRODITEv1101 (Figure 9g,l,q) and PERSIANN_CDR (Figure 9i,n,s) have shown little to no representation of extreme rainfall day counts beyond 50 mm for the whole country, with the latter having a slightly better performance for heavy rainfall days.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

462 *3.5 Statistical error metrics and skill scores*

463 Table 2 shows the RMSE, MAE, R^2 and KS metrics for the four datasets over the daily,
464 monthly and seasonal timescales. In the daily timescale, high RMSE (>10.00 mm day^{-1}) and
465 MAE (>4.00 mm day^{-1}) values were measured for all datasets, with APHRODITEv1101 having
466 the smallest and CHIRPSv2 with the highest values for both. For the monthly and seasonal
467 timescales, the datasets have recorded lower error metrics, which vary from around a tenth to
468 a quarter of the error measured in the daily timescale. For these longer timescales,
469 APRHODITEv1101 still consistently records the lowest MAE; however, TRMM_3B42v7 has
470 the lowest RMSE at monthly and seasonal timescales due to better representation of high
471 rainfall (Figure 4e,j). On the other hand, the largest daily RMSE and MAE were recorded from
472 CHIRPSv2 in the daily scale, and PERSIANN_CDR in the monthly and seasonal scales.

473 APHRODITEv1101 recorded the highest R^2 in the daily scale, but this may be due to a
474 conservative approach of reducing residuals by overrepresenting the most frequently occurring
475 dry and light-to-moderate rainfall days (Figure 4b,g). TRMM_3B42v7 has the next highest R^2
476 in the daily scale, and is tied with APHRODITEv1101 in R^2 for the monthly and seasonal
477 scales. CHIRPSv2 and PERSIANN_CDR have more or less comparable R^2 for all timescales.

478 However, despite consistently obtaining high R^2 for all timescales, APHRODITEv1101
479 has the poorest KS for the daily timescale. This stems from the dataset's percentile distribution
480 that is shifted toward the lower-than-station rainfall values (Figure 5b). On the other hand,
481 CHIRPSv2 and TRMM_3B42v7 have the best KS in the daily scale because these datasets
482 have more closely reproduced the PAGASA rainfall percentile distribution. Lastly, in the
483 longer monthly and seasonal timescales, all datasets have more or less the same performance
484 in terms of KS.

485 The generally meager skill score metrics per daily rain event category for all datasets
486 in the Philippines suggest a strong need for better estimation algorithms and call for added

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

precaution when using these datasets for operational weather forecast verifications (Table 3). As shown by its FBS values, APHRODITEv1101 indicates a tendency to underdetect dry days and largely overdetect light to moderate and heavy rain days, which are consistent with its partitioning of these event categories. However, it has garnered the highest POD values among all datasets, correctly detecting more than 75% to fewer than half of the PAGASA rainfall in the categories from <1 mm to 20-50 mm. Although FAR is quite high, APHRODITEv1101 has the lowest false detections among the datasets, especially for dry days. Also, its PSS suggests that APHRODITEv1101 best discriminates categories from dry days to very heavy rainfall days (50-100 mm). However, despite its relatively better skill metrics, this dataset could be specializing in the more frequent low rainfall events and relies on conservative forecasts for rare high rainfall events, again as evidenced by its partition (Figure 4) and Q-Q plots (Figure 5a).

The FBS values of PERSIANN_CDR has a similar trend with APHRODITEv1101 but with slightly improved detected events for all except the >100 mm category, in which it shows the most propensity to underdetect among all datasets. It has an FBS closest to unity for 20-50 mm among all datasets, but this category was not represented well in the partition plots (Figure 4). It has relatively high POD for dry and light to moderate days but poor POD for higher rainfall days with overall high FAR values across all categories. These and the overall low PSS values indicate the limitations of PERSIANN_CDR in estimating daily rainfall over the country.

TRMM_3B42v7 exhibits the smallest detection bias with FBS values closest to unity across most categories with marginally more frequent underdetection (overdetection) in light to moderate rainfall (dry days). The dataset has a strikingly good detection performance in dry days and in very heavy rainfall events above 50 mm as indicated by high POD values. However, its FAR values are quite high, reaching above 50% for all event categories, except for dry days.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

These values reach even up to 80% for heavy rainfall days, where the POD is also consistently low. It has low PSS for the event categories between 1 to 50 mm, but has high PSS next to APHRODITEv1101 for dry days and 50-100 mm, and more than double the PSS among all datasets for the >100 mm extreme rainfall events.

The FBS values for CHIRPSv2 show that it has the least detection bias for light to moderate rainfall for all datasets, a slight underdetection for dry days, and overdetection for heavy rainfall days. It is the only dataset to have overdetections for the 20-50 mm category, which may well be the reason for its large spread in positive biases. Compared to TRMM, CHIRPSv2 has higher POD in the heavy and 20-50 mm rainfall days but slightly poorer detection for dry days. Its PSS is comparable to TRMM_3B42v7 for all categories, except for a noticeably lower value over the >100 mm category.

4. Conclusions

The performance of four gridded rainfall datasets with respect to PAGASA station rainfall was evaluated using bias analysis, distribution comparisons, and different statistical metrics across different Philippine climate types, event categories, and timescales. Table 4 synthesizes the results through a listing of the potential applications per dataset. APHRODITEv1101 showed an overall propensity to underestimate Philippine rainfall as seen in its consistent dry bias (5.0 to 10.0 mm day^{-1}), underestimated season rainfall peak for climate Type 1 and Type 2 and lack of high rainfall days. It recorded the least errors and desirable skill metrics for non-extreme daily rainfall among all datasets, but this could partly be influenced by an overrepresentation of light rainfall. Thus, APHRODITEv1101 is suited for applications that deal with climatological analysis and non-extreme rainfall, and also those geared on long-term trends since it also has the longest time coverage out of the four datasets.

CHIRPSv2 has the best performance in capturing the seasonal peaks of the prominent climate Type 1 and Type 2. It underestimated rainfall for May to October

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

(8.0 to 10.0 mm day^{-1}) and overestimated for November to February (1.0 to 4.0 mm day^{-1}), but the relatively low bias in monthly and seasonal scales is likely due to the large spread of daily biases. It also has decent representation of rainfall categories and comparably good performance in terms of the statistical error metrics, which makes it a good dataset for computing rainfall-based climate indices. However, it is most prone to larger errors for light rainfall in the daily scale as it overestimates the frequency of heavy rainfall, and it fails to capture rainfall beyond 100 mm. Thus, CHIRPSv2 may be best used in studies concerned at seasonal and climatological scales.

Among the four datasets, PERSIANN_CDR is least able to capture aspects of Philippine rainfall and ranks lowest in almost all metrics. The dataset could be used for climatological analysis because it generally captured the rainfall profile of the climate types. However, it showed a wide range of biases with a predominant high dry bias (10.0-13.0 mm day^{-1}), and poor error metrics in the daily, monthly, and seasonal timescales. Similar deficiencies in the dataset were also observed in studies for other countries (Liu et al. 2015; Dinku et al. 2010)

Lastly, the satellite-derived TRMM_3B42v7 most accurately replicates the frequency and proportion of rain event categories and best represents the spatiotemporal distribution of both dry and torrential rainfall days, making it useful for studies on rainfall distribution, extreme indices, and flood/drought monitoring. It recorded the lowest bias (1.0 to 5.0 mm day^{-1}) and ranks among the highest in terms of statistical metrics in monthly and seasonal timescales. However, it is still subject to high false alarm rates for heavy rainfall and large seasonal biases in certain regions. Nevertheless, this dataset is fit for a wide variety of applications and can be used in investigating climatological and episodic events for the Philippines.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

Acknowledgment

This research is supported by the Department of Science and Technology, Philippines through the Students Support Research Fund (SSRF) of the Accelerated S&T Human Resource Development Program (ASTHRDP) of the Science Education Institute (SEI) and the Philippine Council for Industry, Energy and Emerging Technology Research and Development (PCIEERD), and by the Department of Physics, School of Science and Engineering, Ateneo de Manila University.

References

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeorol.*, **4**, 1147–1167, doi:10.1175/1525-7541(2003).

Ali, G., Q. U. Zaman, G. Ali, G. Rasul, T. Mahmood, Q. Zaman, and S. B. Cheema, 2012: Validation of APHRODITE Precipitation Data for Humid and Sub Humid Regions of Pakistan View project Validation of APHRODITE Precipitation Data for Humid and Sub Humid Regions of Pakistan. *Pakistan Journal of Meteorology*, **9**, 57 pp.

Arulraj, M., and A. P. Barros, 2017: Shallow Precipitation Detection and Classification Using Multifrequency Radar Observations and Model Simulations. *J. Atmos. Ocean. Technol.*, **34**, 1963–1983, doi:10.1175/JTECH-D-17-0060.1.

Ashouri, H., and Coauthors, 2015: PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bull. Am. Meteorol. Soc.*, **96**, 69–83, doi:10.1175/BAMS-D-13-00068.1.

Bagtasa, G., 2017: Contribution of tropical cyclones to rainfall in the Philippines. *J. Clim.*, **30**, 3621–3633, doi:10.1175/JCLI-D-16-0150.1.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

585 Central Intelligence Agency, 2019: East Asia/Southeast Asia::Philippines. *World Factb.*
586 *2019*,. https://www.cia.gov/library/publications/the-world-factbook/geos/rp.html
587 (Accessed November 8, 2019).

588 Cinco, T. A., R. G. de Guzman, F. D. Hilario, and D. M. Wilson, 2014: Long-term trends and
589 extremes in observed daily precipitation and near surface air temperature in the
590 Philippines for the period 1951-2010. *Atmos. Res.*, **145–146**, 12–26,
591 doi:10.1016/j.atmosres.2014.03.025.

592 ——, and Coauthors, 2016: Observed trends and impacts of tropical cyclones in the
593 Philippines. *Int. J. Climatol.*, **36**, 4638–4650, doi:10.1007/s10905-018-9674-0.

594 Coronas, J., 1920: *The Climate and Weather of the Philippines, 1903 – 1918*.

595 Corporal-Lodangco, I. L., and L. M. Leslie, 2017: Defining Philippine Climate Zones Using
596 Surface and High-Resolution Satellite Data. *Procedia Comput. Sci.*, **114**, 324–332,
597 doi:10.1016/J.PROCS.2017.09.068.

598 Cruz, F. T., G. T. Narisma, M. Q. Villafuerte, K. U. Cheng Chua, and L. M. Olaguera, 2013:
599 A climatological analysis of the southwest monsoon rainfall in the Philippines. *Atmos.*
600 *Res.*, **122**, 609–616, doi:10.1016/j.atmosres.2012.06.010.

601 Diem, J. E., J. Hartter, S. J. Ryan, and M. W. Palace, 2014: Validation of Satellite Rainfall
602 Products for Western Uganda. *J. Hydrometeorol.*, **15**, 2030–2038, doi:10.1175/JHM-D-
603 13-0193.1.

604 Dinku, T., F. Ruiz, S. J. Connor, and P. Ceccato, 2010: Validation and intercomparison of
605 satellite rainfall estimates over Colombia. *J. Appl. Meteorol. Climatol.*, **49**, 1004–1014,
606 doi:10.1175/2009JAMC2260.1.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

607 ——, C. Funk, P. Peterson, R. Maidment, T. Tadesse, H. Gadain, and P. Ceccato, 2018:
608 Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Q. J. R.*
609 *Meteorol. Soc.*, **144**, 292–312, doi:10.1002/qj.3244.

610 Flores, J., and V. Balagot, 1969: Climate of the Philippines. Climates of northern and eastern
611 Asia. *Arakawa H World Surv. Climatol. Elsevier, Amsterdam*, **8**, 159–213.

612 Funk, C., and Coauthors, 2015a: The climate hazards infrared precipitation with stations—a
613 new environmental record for monitoring extremes Background & Summary.
614 doi:10.1038/sdata.2015.66.

615 Funk, C., A. Verdin, J. Michaelsen, P. Peterson, D. Pedreros, and G. Husak, 2015b: A global
616 satellite-assisted precipitation climatology. *Earth Syst. Sci. Data*, **7**, 275–287,
617 doi:10.5194/essd-7-275-2015.

618 Ghelli, A., 2009: Verification of categorical predictands. *Int. J. Climatol.*, **34**, 3881–3899,
619 doi:10.1002/joc.3948.

620 Hilario, F., R. De Guzman, D. Ortega, P. Hayman, and B. Alexander, 2009: El Niño Southern
621 Oscillation in the Philippines: Impacts, Forecasts, and Risk Management. *Philipp. J.*
622 *Dev.*, **36**, 9–34.

623 Huffman, G. J., 2017: The Transition in Multi-Satellite Products from TRMM to GPM
624 (TMPA to IMERG). 1–34.

625 Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis
626 (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine
627 Scales. *J. Hydrometeorol.*, **8**, 38–55, doi:10.1175/JHM560.1.

628 Jamandre, C. A., and G. T. Narisma, 2013: Spatio-temporal validation of satellite-based

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

rainfall estimates in the Philippines. *Atmos. Res.*, **122**, 599–608, doi:10.1016/j.atmosres.2012.06.024.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, P. Xie, R. J. Joyce, J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *J. Hydrometeorol.*, **5**, 487–503, doi:10.1175/1525-7541(2004)

Juneng, L., and F. T. Tangang, 2010: Long-term trends of winter monsoon synoptic circulations over the maritime continent: 1962-2007. *Atmos. Sci. Lett.*, **11**, 199–203, doi:10.1002/asl.272.

Kintanar, R., 1984: *Climate of the Philippines, PAGASA Report*. Manila, 38 pp.

Kreft, S., D. Eckstein, L. Junghans, C. Kerestan, and U. Hagen, 2014: *GLOBAL CLIMATE RISK INDEX 2015. Who Suffers Most From Extreme Weather Events? Weather-related Loss Events in 2013 and 1994 to 2013*. Bonn,.

Liu, J., Z. Duan, J. Jiang, and A.-X. Zhu, 2015a: Evaluation of Three Satellite Precipitation Products TRMM 3B42, CMORPH, and PERSIANN over a Subtropical Watershed in China. *Adv. Meteorol.*, **2015**, 1–13, doi:10.1155/2015/151239.

——, ——, ——, and A. X. Zhu, 2015b: Evaluation of three satellite precipitation products TRMM 3B42, CMORPH, and PERSIANN over a subtropical watershed in China. *Adv. Meteorol.*, **2015**, doi:10.1155/2015/151239.

Mason, S. J., 2016: *Guidance on Verification of Operational Seasonal Climate Forecasts*. 79 pp.

Murakami, T., and J. Matsumoto, 1994: Summer Monsoon over the Asian Continent and

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

651 Western North Pacific. *Int. J. Eng. Technol.*, **72**, 719–745,
652 doi:https://doi.org/10.2151/jmsj1965.72.5_719.

653 Prasetia, R., A. R. As-syakur, and T. Osawa, 2013: Validation of TRMM Precipitation Radar
654 satellite data over Indonesian region. *Theor. Appl. Climatol.*, **112**, 575–587,
655 doi:10.1007/s00704-012-0756-1.

656 Pullen, J., A. L. Gordon, M. Flatau, J. D. Doyle, C. Villanoy, and O. Cabrera, 2015:
657 Multiscale influences on extreme winter rainfall in the Philippines. *J. Geophys. Res.*,
658 **120**, 3292–3309, doi:10.1002/2014JD022645.

659 dos Reis, J., C. Rennó, E. Lopes, J. B. C. Dos Reis, C. D. Rennó, and E. S. S. Lopes, 2017:
660 Validation of Satellite Rainfall Products over a Mountainous Watershed in a Humid
661 Subtropical Climate Region of Brazil. *Remote Sens.*, **9**, 1240, doi:10.3390/rs9121240.

662 Rudolf, B., 1993: Management and analysis of precipitation data on a routine basis.
663 *Proceedings of International Symposium on Precipitation and Evaporation*, Vol. 1 of,
664 69–76.

665 Usman, M., J. E. Nichol, A. T. Ibrahim, and L. F. Buba, 2018: Agricultural and Forest
666 Meteorology A spatio-temporal analysis of trends in rainfall from long term satellite
667 rainfall products in the Sudano Sahelian zone of Nigeria. *Agric. For. Meteorol.*, **260**–
668 **261**, 273–286, doi:10.1016/j.agrformet.2018.06.016.

669 Villafuerte, M. Q., J. Matsumoto, I. Akasaka, H. G. Takahashi, H. Kubota, and T. A. Cinco,
670 2014: Long-term trends and variability of rainfall extremes in the Philippines. *Atmos.*
671 *Res.*, **137**, 1–13, doi:10.1016/j.atmosres.2013.09.021.

672 World Meteorological Organization, 2018: *Guidelines on the Defintion and Monitoring of*

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

673 *Extreme Weather and Climate Events*. 62 pp.

674 Xie, P., B. Rudolf, U. Schneider, and P. A. Arkin, 1996: Gauge-based monthly analysis of
675 global land precipitation from 1971 to 1994. *J. Geophys. Res. Atmos.*, **101**, 19023–
676 19034, doi:10.1029/96jd01553.

677 Yatagai, A., and Coauthors, 2012: APHRODITE: Constructing a Long-Term Daily Gridded
678 Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bull. Am.*
679 *Meteorol. Soc.*, **93**, 1401–1415, doi:10.1175/BAMS-D-11-00122.1.

680 Zhang, H., T. Casey, H. Zhang, and T. Casey, 2000: Verification of Categorical Probability
681 Forecasts. *Weather Forecast.*, **15**, 80–89, doi:10.1175/1520-0434(2000)

682

683

684

685

686

687

688

689

690

691

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

692 TABLE 1. Daily rainfall categories as defined in this study

Event Category	Rainfall range
Dry days	< 1mm
Light to moderate rain days	1 to 10 mm
Heavy rain days	10 to 20 mm
Very heavy rain days	> 20 mm

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

TABLE 2. RMSE, MAE, R^2 and KS of the gridded rainfall datasets averaged for all stations for the daily, monthly and seasonal timescales. Values in bold indicate the metric that is closest to ideal among the datasets.

Timescale	Metric	APHRODITEv1101	CHIRPSv2	PERSIANN_CDR	TRMM_3B42v7
day	RMSE	**11.3**	16.4	15.9	15.6
	MAE	**4.5**	7.4	7.2	6.6
	R^2	**0.6**	0.3	0.3	0.4
	KS	0.4	**0.1**	0.2	**0.1**
month	RMSE	3.0	3.4	4.2	**2.8**
	MAE	**1.9**	2.3	2.9	2.0
	R^2	**0.8**	0.7	0.7	**0.8**
	KS	**0.2**	**0.2**	**0.2**	**0.2**
season	RMSE	2.1	2.1	3.0	**2.0**
	MAE	**1.5**	1.6	2.3	1.6
	R^2	**0.9**	0.8	0.7	**0.9**
	KS	**0.2**	**0.2**	0.3	**0.2**

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

TABLE 3. Frequency Bias Score, POD, FAR and PSS of the gridded rainfall datasets for the different rainfall event categories. Values in bold indicate the metric that is closest to ideal among the datasets.

Skill Metric	Event Category	APHRODITEv1101	CHIRPSv2	PERSIANN_CDR	TRMM_3B42v7
FBS	<1mm	0.65	0.91	0.69	**1.02**
	1-10mm	2.01	**1.01**	1.69	0.91
	10-20mm	1.37	1.51	1.70	**1.03**
	20-50mm	0.82	1.19	**1.00**	0.98
	50-100mm	0.48	0.86	0.50	**0.95**
	>100mm	0.31	0.53	0.12	**1.10**
POD	<1mm	0.62	0.72	0.57	**0.80**
	1-10mm	**0.78**	0.31	0.47	0.31
	10-20mm	**0.44**	0.23	0.25	0.18
	20-50mm	**0.41**	0.30	0.25	0.28
	50-100mm	**0.23**	0.20	0.13	**0.23**
	>100mm	0.27	0.17	0.04	**0.46**
FAR	<1mm	**0.05**	0.21	0.18	0.21
	1-10mm	**0.61**	0.69	0.72	0.66
	10-20mm	**0.68**	0.85	0.85	0.82
	20-50mm	**0.50**	0.75	0.75	0.71
	50-100mm	**0.52**	0.76	0.75	0.75
	>100mm	**0.12**	0.68	0.63	0.58
PSS	<1mm	**0.57**	0.42	0.38	0.46
	1-10mm	**0.42**	0.10	0.11	0.14
	10-20mm	**0.36**	0.13	0.13	0.11
	20-50mm	**0.38**	0.23	0.19	0.23
	50-100mm	**0.22**	0.19	0.12	**0.22**
	>100mm	0.27	0.17	0.04	**0.45**

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

762 TABLE 4. Potential Philippine climate applications of each gridded rainfall dataset

763

Potential Applications	APHRODITEv1101	CHIRPSv2	PERSIANN_CDR	TRMM_3B42v7
Seasonal analysis *good seasonal error metrics*		✓		✓
Climatological analysis *captures climate type rainfall profiles*	✓	✓	✓	✓
Rainfall distributions *captures rainfall event partitions and contributions*				✓
Rainfall-based climate indices *good rainfall distribution, good error metrics for all temporal scales*		✓		✓
Drought monitoring *good skill metrics and spatial distribution for rainfall less than 1.0 mm*				✓
Flood monitoring *good skill metrics and spatial distribution for rainfall greater than 50.0 mm*				✓
Non-extreme daily rainfall *good skill metrics and spatial distribution for rainfall within 1.0-50.0 mm*	✓			

764

765

766

767

768

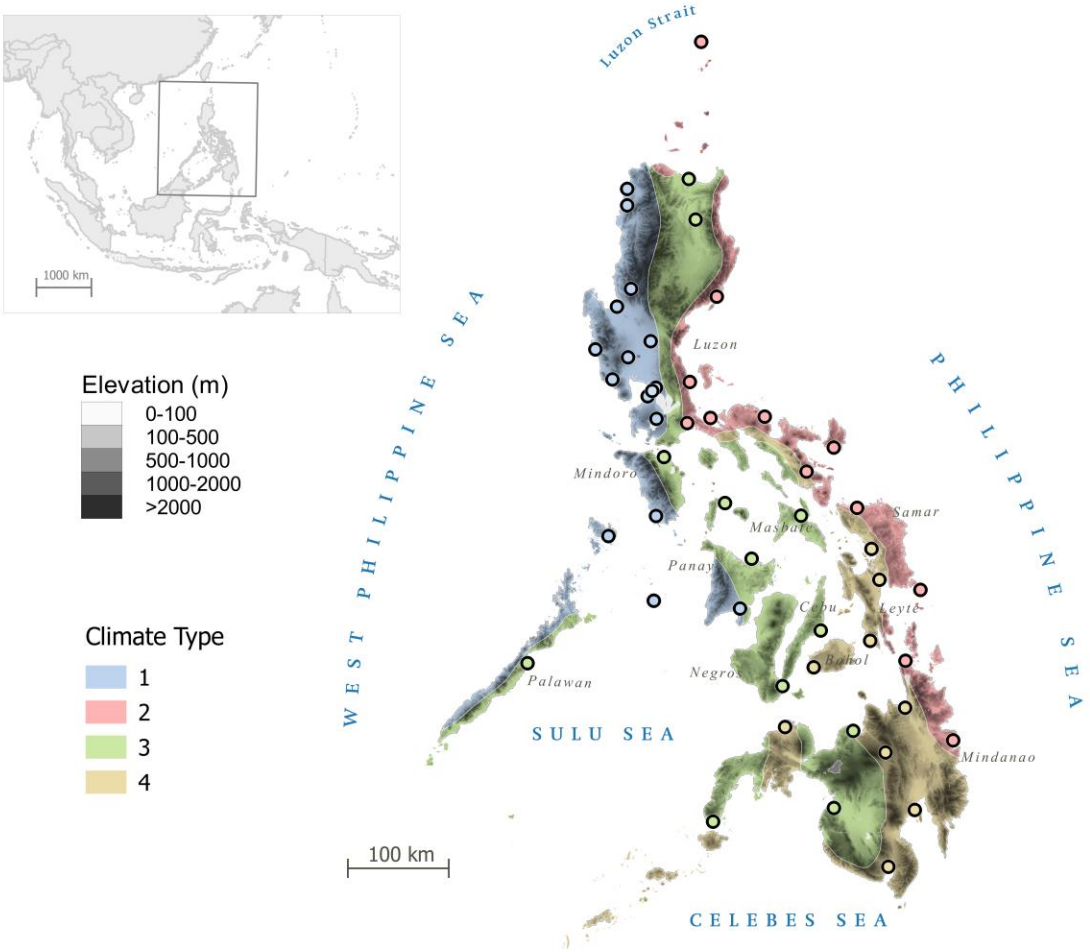
769

770

771

772

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



773

774 FIGURE 1. Map of the climate types of the Philippines and location of the 49 PAGASA
775 stations used. Elevation is indicated as shading. This map does not show the full extent of the
776 maritime borders of the country.

777

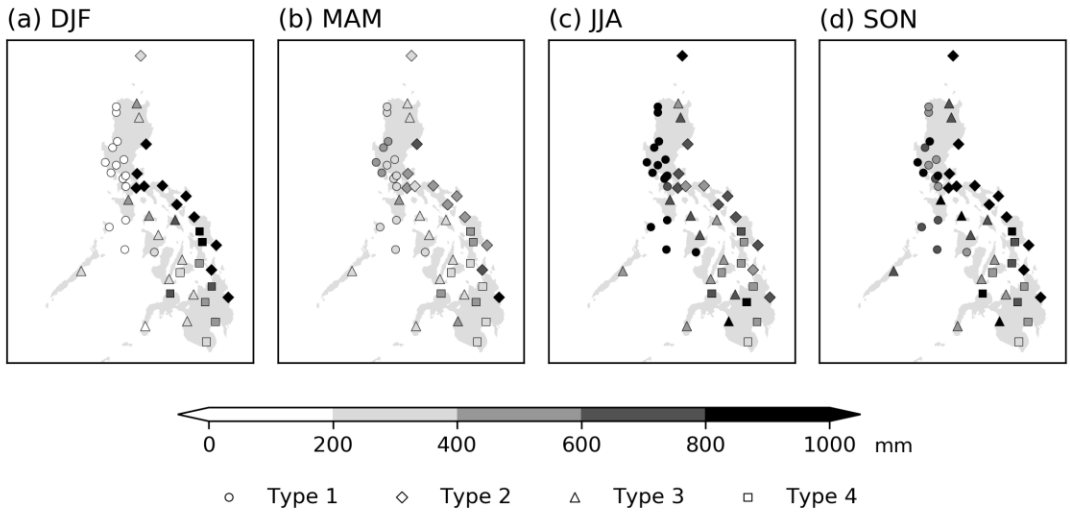
778

779

780

781

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



782

783 FIGURE 2. Average seasonal total rainfall from PAGASA stations for 1998-2005

784

785

786

787

788

789

790

791

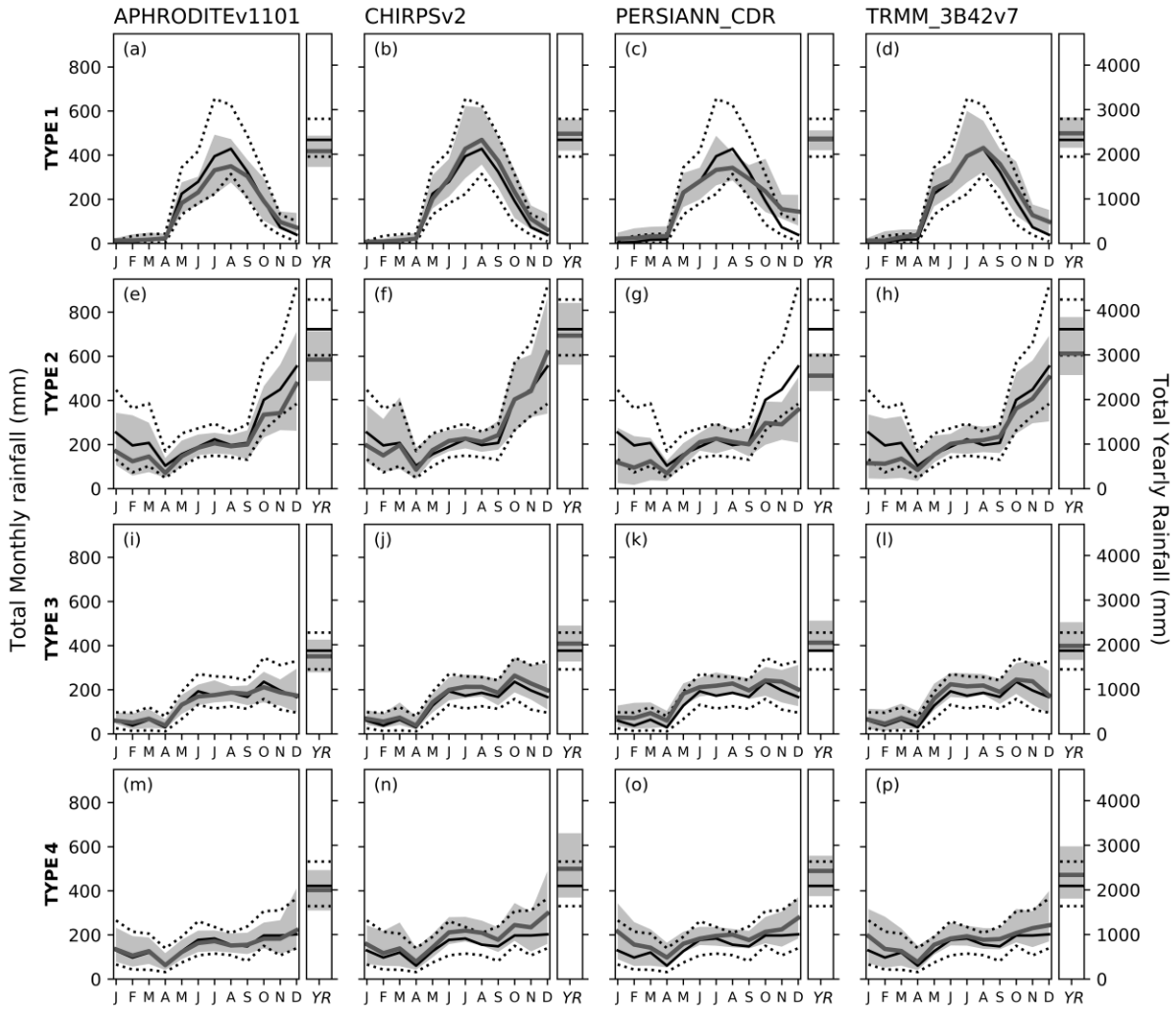
792

793

794

795

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



796

797 FIGURE 3. Monthly and annual rainfall median and spread of PAGASA and the four datasets
798 for each climate type. Solid black lines represent the median total rainfall for PAGASA while
799 the solid gray lines are for each of the datasets. The upper and lower dotted lines represent the
800 PAGASA P75 and P25 respectively, while the gray shadow represents the P25 to P75 range
801 of the datasets.

802

803

804

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

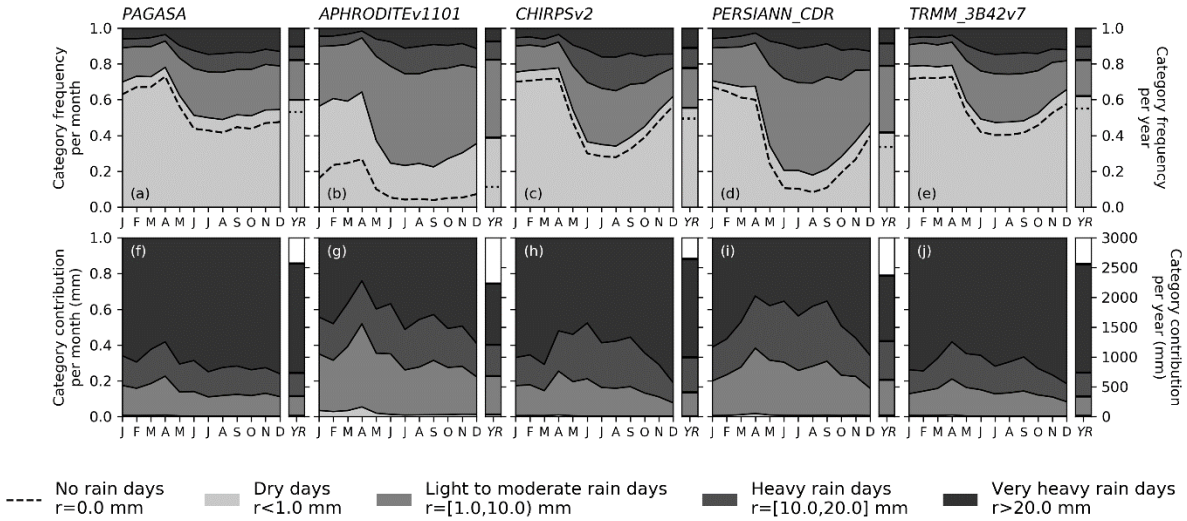
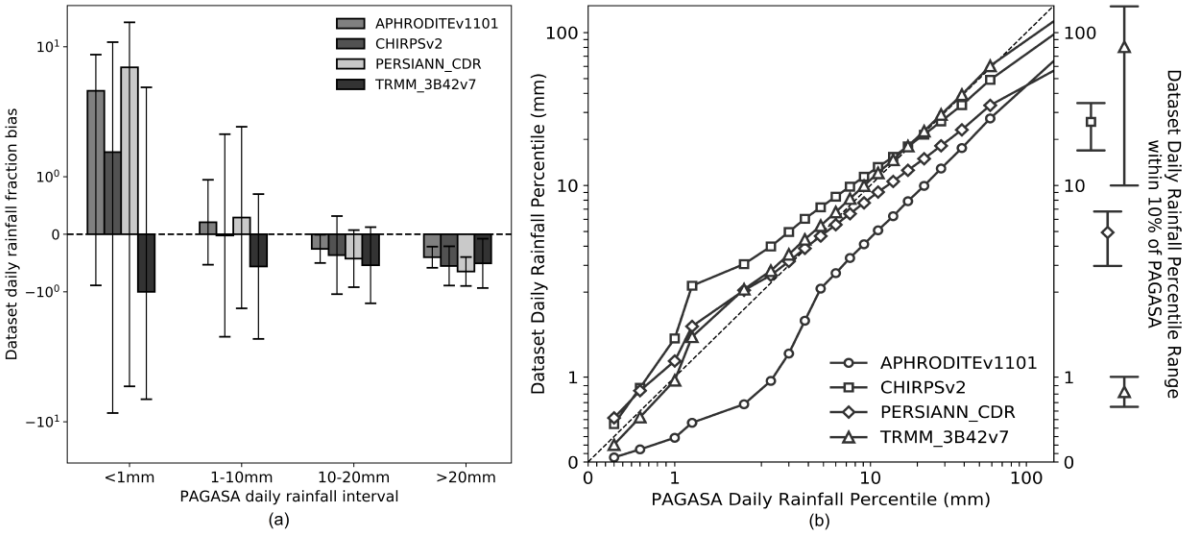


FIGURE 4. Partition plots of the monthly and annual frequency (top row) and contribution (bottom row) of the four rain event categories recorded in PAGASA and the four datasets over 1998-2005. No rain days (r=0.0mm), a subset of dry days, are represented by a dashed line in the frequency plots.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



819

820 FIGURE 5. (a) Daily rainfall bias distribution of the datasets as fraction of the PAGASA daily
821 rainfall per rain event category. The bars denote the median while the error bars indicate the
822 P25 and P75 fraction bias. (b) Q-Q plot of the four datasets excluding no rain days. Points are
823 calculated with an interval of one percentile. The vertical lines at the right side of the Q-Q plot
824 indicate the range where the rainfall percentile of each dataset fell within 10% of PAGASA.

825

826

827

828

829

830

831

832

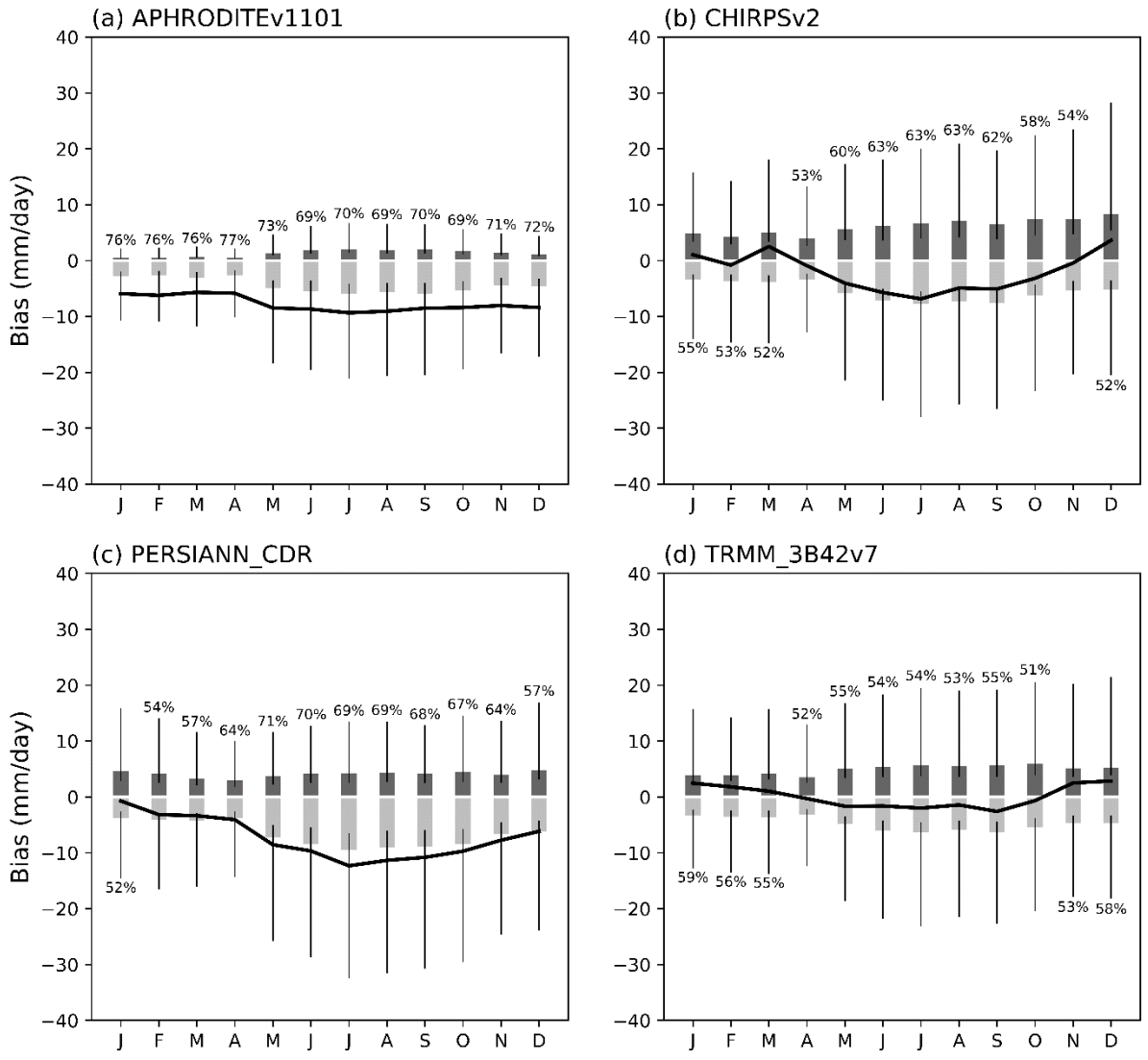
833

834

835

836

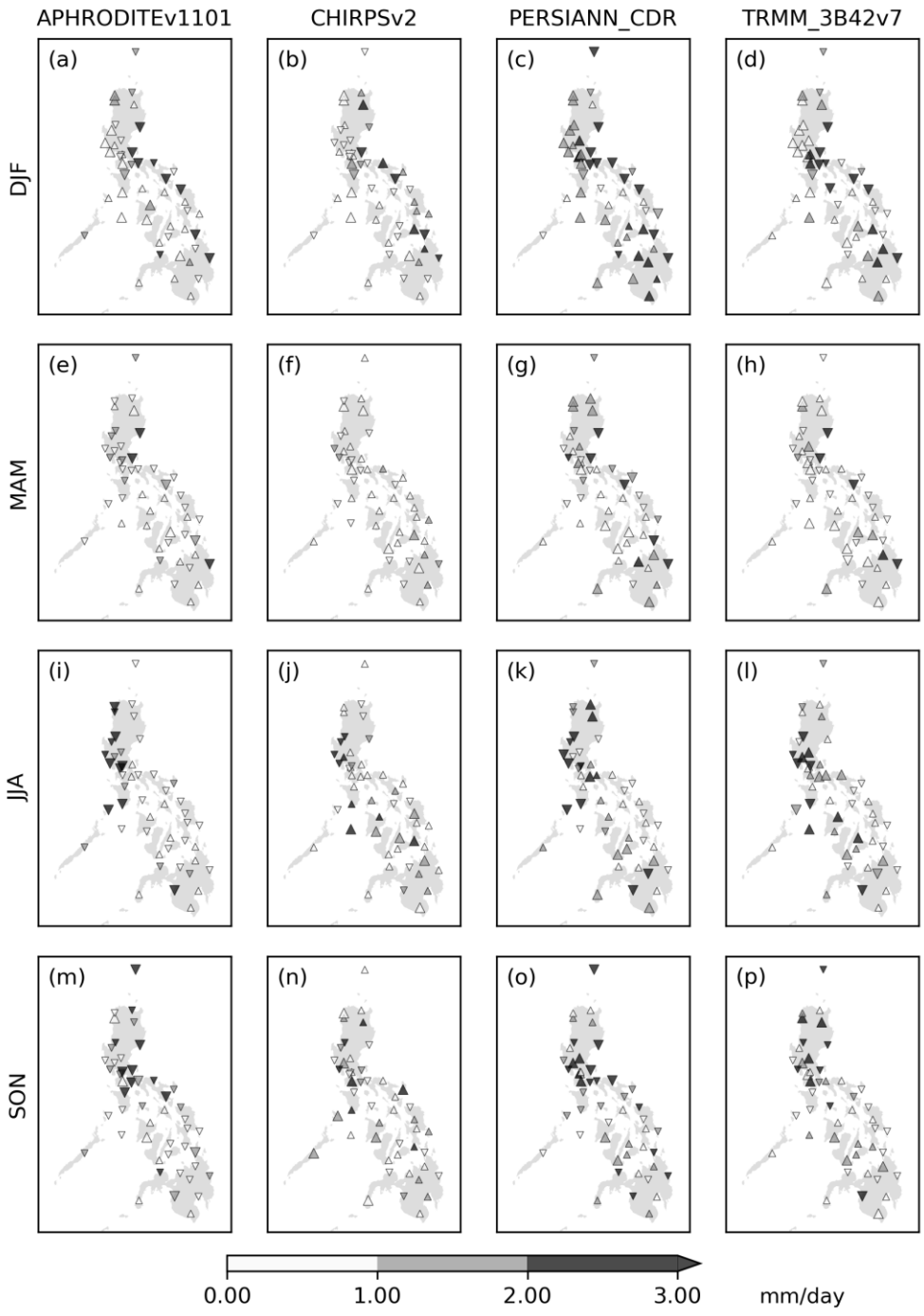
Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



837

838 FIGURE 6. Bidirectional box plots of the bias of the gridded datasets with respect to PAGASA.
839 The thick black line indicates the monthly mean bias. The dark gray and light gray bars denote
840 the median daily positive and negative bias per month, respectively, while the endpoints of
841 vertical error bars denote the P25 and P75 of each bias sign per month. The percentages located
842 above/below the error bars indicate the bias sign that occurred most frequently for that
843 particular month.

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.

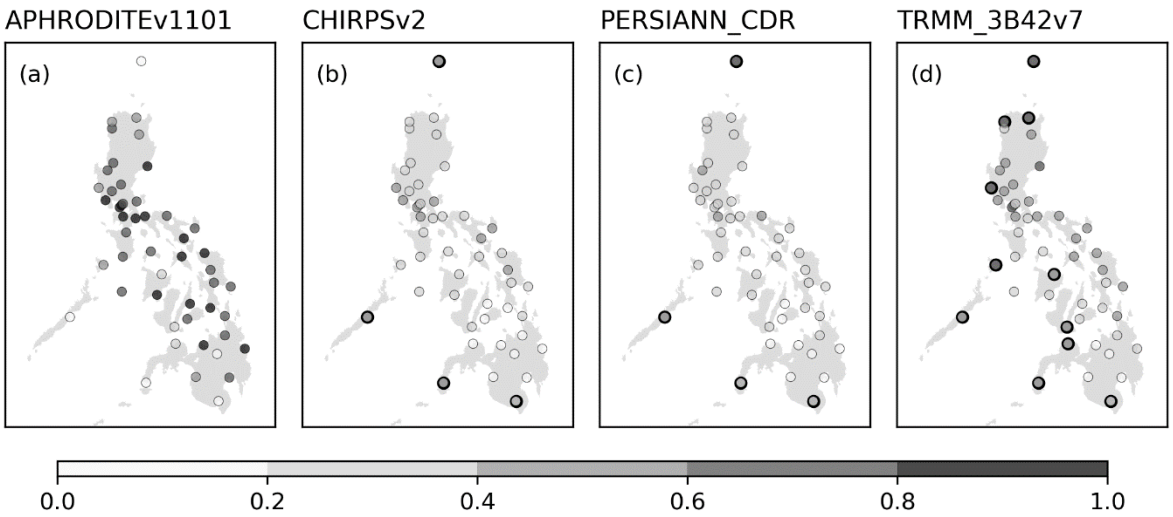


844

845 Figure 7. Seasonal mean daily bias map of the four datasets with respect to PAGASA for 1998-
846 2005. Stations where wet (dry) bias was recorded is marked by a triangle pointing up (down).
847 Also, stations whose individual monthly bias signs are consistent with the seasonal mean bias
848 sign for more than half of all months within the season are marked with a larger symbol.

849

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



850

851 FIGURE 8. Daily correlation map of the four datasets with respect to PAGASA for 1998-2005.
852 Markers with heavier outline in (b) to (d) indicate locations where satellite-derived datasets
853 have higher correlation values than station-based dataset APHRODITEv1101.

854

855

856

857

858

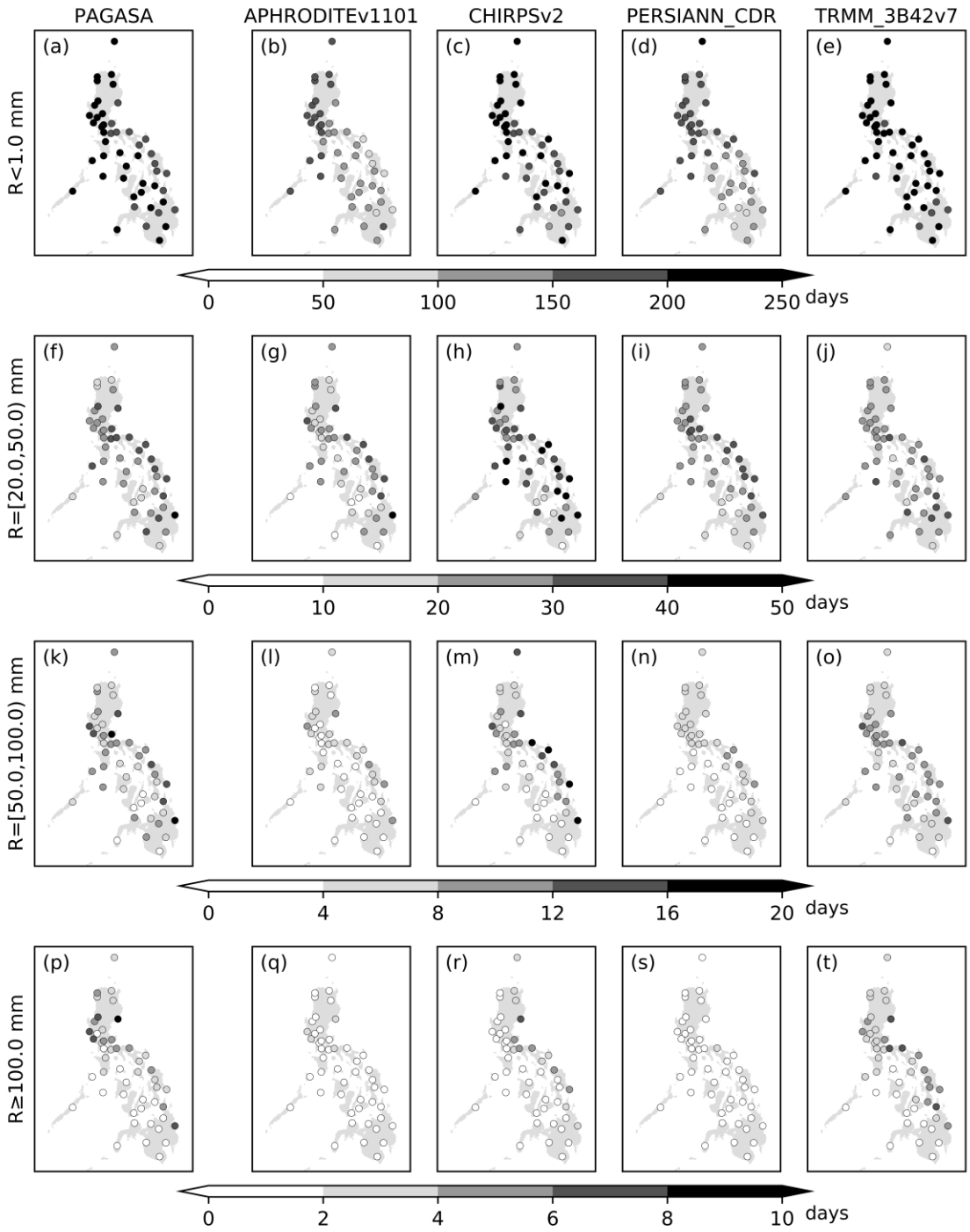
859

860

861

862

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.



863

864 FIGURE 9. Spatial distribution of the frequency of each event category of PAGASA and the
865 four datasets for 1998-2005.

866

Accepted for publication in *Journal of Hydrometeorology*. DOI 10.1175/JHM-D-19-0276.1.