

Regional-scale evaluation of 14 satellite-based precipitation products in characterising extreme events and delineating rainfall thresholds for flood hazards

Geraldo Moura Ramos Filho[a,*], Victor Hugo Rabelo Coelho[b], Emerson da Silva Freitas[a,c], Yunqing Xuan[d], Luca Brocca[e], Cristiano das Neves Almeida[a]

[a] Department of Civil and Environmental Engineering, Federal University of Paraíba, João Pessoa, 58051-900, Brazil

[b] Department of Geosciences, Federal University of Paraíba, João Pessoa, 58051-900, Brazil

[c] Federal Institute of Education, Science and Technology of Pernambuco, Pesqueira, 55200-000, Brazil

[d] College of Engineering, Swansea University, Bay Campus, Swansea, SA1 8EN, UK

[e] National Research Council, Research Institute for Geo-Hydrological Protection, Perugia, 06128, Italy

***** Corresponding author. Tel.: +558332167684. E-mails: Geraldo M. Ramos Filho (geraldo.ramos@academico.ufpb.br) · Victor Hugo R. Coelho (victor.coelho@academico.ufpb.br) · Emerson da S. Freitas (emerson.sfreitas@hotmail.com) · Yunqing Xuan (y.xuan@swansea.ac.uk), Luca Brocca (luca.brocca@irpi.cnr.it), Cristiano das N. Almeida (cristiano.almeida@academico.ufpb.br)

Geraldo M. Ramos Filho (https://orcid.org/0000-0002-7819-2511) · Victor Hugo R. Coelho (https://orcid.org/0000-0002-6669-4856) · Emerson da S. Freitas (https://orcid.org/0000-0002-8409-8206) · Yunqing Xuan (https://orcid.org/0000-0003-2736-8625) · Luca Brocca (https://orcid.org/0000-0002-9080-260X) · Cristiano das N. Almeida (https://orcid.org/0000-0002-3655-5396)

Abstract: Gridded satellite-based rainfall products have not been so far evaluated for flood hazards monitoring through empirical methods, especially over large areas. Therefore, the main aims of this study are (i) to assess the quality of satellite-based precipitation products for identifying extreme rainfall events able to produce flood events, and (ii) to evaluate the use of satellite-based precipitation products for creating rainfall thresholds to support decision making for issuing flood warnings. Eight products fully based on satellite data (i.e., uncorrected) and six gauge-corrected products were evaluated based on ground-based data obtained from 583 sub-daily rainfall gauges and considering a catalogue of 551 flood occurrences in the state of São Paulo, Brazil, for a period of five years (2015-2019). The gauged values were compared with the precipitation products for six different time steps (3 h, 6 h, 12 h, 1 d, 3 d, and 10 d) and considering rainfall-duration thresholds for six non-exceedance probabilities (5%, 10%, 20%, 30%, 40%, and 50%). Results show that all analysed products tend to large underestimate the extreme rainfall events (i.e., when and where flood events were registered), mainly for sub-daily scales, with the best results found for two uncorrected products (i.e., PDIR-Now and GPM+SM2RAIN) considering 10-days accumulated precipitation. Considerable underestimations were also identified for the rainfall thresholds delineated by the satellite-based products, with the best performances obtained by CHIRP V2.0 (uncorrected) and IMERG-F (corrected) considering tolerance levels of 20%. Based on our findings, the rainfall satellite-based products dataset, even less accurate than the ground-based observations, can be applied, when multi-daily accumulated data are considered, as an alternative source of data for determining precipitation thresholds in some regions that present low-density of rain gauges but not replace the gauged data in regions with a high-density of rain gauges with sub-daily data available.

Keywords: extreme rainfall events, empirical rainfall thresholds, flood hazards.

1. Introduction

Floods occur every year in almost all countries, causing thousands of deaths, considerable structural damages, and significant economic losses worldwide (Dinis et al., 2021; Hallegatte et al., 2013; Sampson et al., 2015). According to the United Nations Office for Disaster Risk Reduction (UNDRR), the number of flood occurrences increased 2.3 times between 2000 and 2019 compared to the previous twenty years (i.e., from 1980 to 1999). This growing number of flood events over the years is mainly attributed to the extreme weather conditions, high urbanisation rate, and inadequate response to disasters (Du et al., 2015; Špitalar et al., 2014; Tsakiris, 2014). These flood events accounted for approximately 44% of all-natural disasters that occurred from 2000 to 2019 and affected more than 1.5 million people in almost all countries in the world, causing more than 104,600 deaths and US$ 651 billion of economic losses (UNDRR and CRED 2020).

Great efforts have been made during the last few decades to develop and improve methods for a better prediction and warning of flood events (Getirana et al., 2020; Young et al., 2021). Such methods for predicting and warning of hydrological disasters are aimed to reduce the damages and deaths caused by floods(Froidevaux et al., 2015; Jang, 2015). Complex computer models, such as hydrodynamic models, are among the most widely used tools to simulate detailed flood dynamics. These hydrodynamic models are undergoing progress both in accuracy and computational efficiency, however, they require a set of detailed input data or a high computation cost (Teng et al., 2017). Therefore, empirical methods still prevail as an alternative approach for flood monitoring (Ramos Filho et al., 2021; Yang et al., 2016), especially in regions where the detailed input data, used to run the aforementioned hydrodynamic models, are scarce or unavailable. For these regions, the rainfall threshold approaches represent a popular tool used to study the relationship between rainfall and hydrological disasters (e.g., floods, flash floods, debris flows and landslides) (Aleotti, 2004; Berti et al., 2012; Glade et al., 2000; Mirus et al., 2018; Santos and Fragoso, 2016; Scheevel et al., 2017).

To support the decision-making processes, several rainfall threshold-based approaches have been developed and they commonly employ two thresholds (upper and lower) that are determined by the properties derived from rainfall events (e.g., intensity, duration, antecedent precipitation), to define the rainfall conditions that are likely to trigger flood events (Diakakis, 2012; Papagiannaki et al., 2015). Recently, the study carried out by Ramos Filho et al. (2021) improved the rainfall threshold identification process that reduces the uncertainties and minimises the number of false alarms for issued flood events. The same study also observed that a considerable amount of flood information could not be used to create and improve further the rainfall threshold method due to the low quality of the observation data and/or the low density of the rain gauges in some areas of São Paulo State, in Brazil.

The use of accurate and spatially well-distributed sub-daily rainfall data, alongside the knowledge about its properties (e.g., depth, duration, intensity, frequency, dry time), is recognised by the scientific community as an essential step to create robust hydrological disaster early warning systems (Chikoore et al., 2021; Dunkerley, 2019; Shrestha et al., 2019). However, obtaining sub-daily rainfall data over large areas from in-situ observations is still a hard task because this type of data records sparsely only covers the global landmass (Hegerl et al., 2015; Lewis et al., 2019). The number of in-situ sub-daily rainfall records is even lower for tropical regions, probably due to the higher installation and operation costs for sub-daily measurements than those at daily timescale (Freitas et al., 2020; Hegerl et al., 2015; Kidd et al., 2017). For instance, Blenkinsop et al. (2018) identified that countries from Africa and Latin America have the lowest availability of sub-daily rainfall data. Consequently, empirical methods that rely on the use of sub-daily rainfall data cannot be properly applied in these data-sparse areas.

The cutting-edge satellite-borne remote sensing technology has played a key role over the recent decades in providing sub-daily rainfall data (Levizzani et al., 2018; Sungmin and Kirstetter, 2018; Tan et al., 2014). Currently, a plethora of promising recently released and revised gridded satellite-based products, providing valuable distributed information of sub-daily rainfall data, are available to be used for many applications (Llauca et al., 2021; Yuan et al., 2019). The characteristics of these remote sensing rainfall products differ in spatial and temporal resolutions (from 0.04º to 2.5º and from 30 minutes to monthly, respectively), spatial coverage (from continental to fully global), and latency (from 15 minutes to several years), among others (Beck et al., 2017b). During the last few decades, several studies have assessed the accuracy of one or a set of satellite-based rainfall data at various spatial and temporal scales, most of which from independent gauge or radar observations (e.g., Tan and Duan, 2017; Gadelha et al., 2018; Wang et al., 2018; Beck et al., 2019a). Some of these studies evaluated the performance of the satellite-based rainfall products regionally or globally for some specific hydrological applications, such as water resources management (e.g., Ranghetti et al., 2018; Sheffield et al., 2018), groundwater storage and depletion (e.g., Vasco et al., 2019; Singh and Saravanan, 2020), hazard monitoring (e.g., Pandey and Srivastava, 2019; Parker et al., 2021), and streamflow modelling (e.g., Su et al., 2019; Camici et al., 2020; Kha et al., 2020; Almagro et al., 2021). However, applications of satellite-based rainfall data for hydrological disasters warning purposes through the use of empirical methods have been scarce, mainly, because: 1) the bias in near real-time rainfall estimates, 2) the latency of products, and 3) insufficient spatial and temporal resolutions (AghaKouchak et al., 2015; Brocca et al., 2017). Based on a literature review, we identified that only a few studies evaluated the capability of the satellite gridded rainfall datasets in detecting landslide events with the use of empirical rainfall thresholds (e.g., Nanda Pratama et al., 2017; Brunetti et al., 2018, 2021; Monsieurs et al., 2019; Chikalamo et al., 2020; He et al., 2020), with no similar analysis for flood events, especially over large areas. A study carried out by Brunetti et al. (2018), for instance, showed that the four analysed precipitation satellite-based products were able to identify landslides occurrences in Italy by adjusting the rainfall thresholds, but with less accuracy than ground-based rainfall observations. More recently, a study performed by Brunetti et al. (2021) in India, also using empirical rainfall thresholds derived from the analysis of historical landslide events, found that the two analysed satellite-based rainfall products outperformed the ground observations thanks to their better spatial and temporal resolutions.

Clearly, satellite-based data are an important data source for improving the spatial representativeness of rainfall-threshold approaches and, consequently, providing tools to create more robust warning systems for flood occurrences, especially in many parts of the world with low-density sub-daily rain gauge networks. Therefore, we commissioned this study to addresses the following scientific questions: (a) How do the currently available rainfall satellite-based products perform for flood events detection? (b) Which satellite-based product performs better in defining empirical rainfall-threshold methods for floods?

The main aims of this study are: (i) to assess the quality of satellite-based precipitation products for identifying extreme rainfall events able to produce flood events, and (ii) to evaluate the use of satellite-based precipitation products to create rainfall thresholds for flood hazards. To achieve the proposed objectives, we used detailed information on flood occurrences available for the São Paulo State in Brazil for a period of five years (2015-2019). In addition, we used a ground-based sub-daily rainfall dataset obtained from a network of around 730 rain gauges and 14 satellite-based precipitation products with different temporal and spatial resolutions. This study is intended then to provide a valuable tool for flood warning systems using satellite-based rainfall products in tropical regions.

2. Study area

This study is the Brazilian state of São Paulo, which has an area of 248,200 km^2 and is located between 19°55'58"S-25°00'53"S and 50°32'15"W-47°55'36"W (Fig. 1). São Paulo is the most populated state of Brazil with approximately 46.6 M inhabitants (IBGE, 2021). According to Alvares et al. (2013), the state has two Köppen’s climate zones (tropical and humid subtropical). The tropical climate zone has a mean annual air temperature above 22 °C and an average annual rainfall above 2,000 mm. Meanwhile, the humid subtropical climate has a mean annual air temperature of 20 °C and an average annual rainfall equal to 1,400 mm $year^{-1}$. The rainfall in this state is more concentrated during the austral spring-summer (i.e., from October to March).

INSERT FIG. 1 HERE

Fig. 1. (a) Map of the São Paulo State showing (a) the 583 rain gauges with the elevation for the state of São Paulo and (b) the 551 flood occurrences with the Köppen’s classification map according to Alvares et al. (2013).

São Paulo State is a global hotspot frequented by many hydrological disaster problems arising from prolonged or intense rainfall events (e.g., landslides, soil erosion, floods, and flash floods), mainly because the natural characteristics of the region, associated with the high level of urbanisation (Tominaga et al., 2015). From 2000 to 2015, the number of natural disasters recorded in São Paulo surpassed 10,800, causing 534 deaths and affecting more than 971,500 people (Brollo and Ferreira, 2016).

3. Materials and methods

3.1 Flood dataset

Information of floods that occurred between January 2015 to December 2019 in the São Paulo State was obtained from the following four sources: (1) The Integrated Storm Monitoring, Forecasting and Alerting System for the Brazilian South-Southeast Regions (SIMPAT); (2) the Brazilian National Centre for Monitoring Early Warning of Natural Disasters (CEMADEN); (3) The Civil Defence of São Paulo State; and (4) press news. Overall, the information obtained from these four sources includes: the disaster type, location (addresses or coordinates), day of occurrence, and the number of affected people, including deaths. Therefore, the information about the extension of the floods was not available in these data sources (i.e., only punctual information), which makes challenging a more complex analysis considering, for instance, many gauges/pixels in a river basin scale. Moreover, the occurrences with (1) daily rainfall less than 10 mm registered near to the flood events or (2) the nearest rain gauge located more than 20 km from the flood events were excluded for further analyses. We identified 762 occurrences of floods in the São Paulo State during the studied period. After the restrictions mentioned above, a total of 551 occurrences of floods were used for further analyses (Fig. 1a). The mean distance between the occurrences and the nearest rain gauge was ~7 km. Most of the 211 flood events were excluded from the analyses because of the lack of rain gauges distant less than 20 km from the occurrences. This exclusion is a consequence of the uneven distribution of rain gauges over the region, which is more concentrated in larger cities.

3.2 Gauged rainfall dataset

This study began by considering ground-based sub-daily rainfall data from 730 automated rain gauges operated by CEMADEN over the period between January 2015 and December 2019. CEMADEN has a national-wide ground-based rainfall network consisting of tipping bucket gauges with a 10-min temporal resolution when it rains and 60-min over no-rain periods. The gauged rainfall data used in this study underwent the same quality control measure as that used by Freitas et al. (2020) to detect possible rain gauge inconsistencies and select only high-quality data. Therefore, only those rain gauges with less than 30 days of missing data along each civil year were considered in this study. Moreover, the gauges that met this criterion were visually inspected as follows: 1) comparing the monthly and sub-daily rainfall data with the five nearest stations to verify large discrepancies between them, and 2) analysing the range of values and changes over subsequent measurements of each rain gauge to identify constant or null rainfall records that probably indicate gauge clogging. After the quality control procedure adopted in this study, a total of 583 gauges were selected to define the rainfall events and calculate their respective rainfall thresholds (Fig. 1b).

3.3 Satellite-based rainfall products

The performance of the 14 (sub-) daily satellite-based rainfall products was evaluated in this study based on a point-to-cell analysis comparison between these estimated datasets and the rain gauges. Table 1 provides objective-focused tabular information of all estimated rainfall datasets considered in this evaluation. All analysed products are global or quasi-global, with data available to cover the entire study period, except the GPM+SM2RAIN that provides rainfall data only until 2018. The spatial resolution of the evaluated rainfall products ranges from 0.04º to 0.5º, while the temporal resolution varies between 30-min and daily. Among the 14 rainfall products, eight of them are fully based on satellite data (hereafter referred to as the uncorrected products, which includes CHIRP V2.0 (Funk et al., 2015), IMERG-E (Huffman et al., 2019), IMERG-L (Huffman et al., 2019), PDIR-Now (Nguyen et al., 2020), PERSIANN (Sorooshian et al., 2000), PERSIANN-CCS (Hong et al., 2004), SM2RAIN-ASCAT V1.2 (Brocca et al., 2019), and GPM+SM2RAIN (Massari et al., 2020)) and six products combine gauge and satellite data (hereafter corrected products, which includes CMORPH-CRT V1.0 (Joyce et al., 2004; Xie et al., 2017), IMERG-F (Huffman et al., 2019), PERSIANN-CDR V1R1 (Ashouri et al., 2015), CHIRPS V2.0 (Funk et al., 2015), MERRA-2 (Gelaro et al., 2017), and MSWEP V2.2 (Beck et al., 2019b, 2017a)). Some of them also use (re)analysis data to generate the products (e.g., CHIRP V2.0, CHIRPS V2.0, MERRA-2, and MSWEP V2.2). A rainfall depth threshold of 0.1 mm day^{-1} was established to define rain/no-rain and to exclude daily events deemed insignificant, following Li and Liu (2020).

Table 1. Summary of 14 precipitation estimates products evaluated in this study, similar to as presented by Beck et al. (2019a). The symbols * and ** highlight the uncorrected (solely satellite data) and corrected (satellite and gauge data) products, respectively.

INSERT TABLE 1 HERE

3.4 Rainfall events and threshold definition

The delineation of the thresholds was based on an empirical model that evaluates the amounts of precipitation that may or may not lead to flooding events through the analyses of the exceedance or not of a certain threshold. Six aggregation periods (3 h, 6 h, 12 h, 1 d, 3 d, and 10 d) were considered to determine the accumulated precipitation. For the sub-daily aggregations, we considered the maximum moving sum of the accumulated precipitation. The daily precipitation was classified according to CEMADEN into light rain (< 10 mm), moderate rain (≥ 10 mm and < 30 mm), heavy rain (≥ 30 mm and < 70 mm), and severe rain (≥ 70 mm).

The rainfall thresholds were determined for the gauged data and each satellite-based rainfall product separately by applying the adapted empirical methodology used by Diakakis (2012) and Papagiannaki et al. (2015). Specifically in this study, we used the accumulated rainfall-duration thresholds for detecting the occurrence of floods, i.e., by plotting the cumulated rainfall of various time intervals against their respective durations. An analysis of the graph based on the following three criteria was performed to define the time interval of the cumulated rainfall that better represents the flood events: (1) the higher number of occurrences above the threshold, (2) the higher amount of non-occurrence below the threshold, and (3) the values of the metrics presented in the next section of this manuscript. Multiple rainfall-duration thresholds were defined from the application of 5%, 10%, 20%, 30%, 40%, and 50% non-exceedance probability aiming to reduce the uncertainties of false alarms.

3.5 Comparison and evaluation procedures

The first evaluation step was to apply the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) scores to assess the performance of the satellite-based rainfall products in characterising the rainfall events that are able to trigger floods. KGE is an objective performance metric combining correlation (CC, represented by the Pearson's correlation coefficient), BIAS (represented by the ratio of estimated and observed means), and variability (VAR, represented by the ratio of the estimated and observed coefficients of variation):

$$\mathrm{KGE} = 1 - \sqrt{(\mathrm{CC}-1)^2 + (\mathrm{BIAS}-1)^2 + (\mathrm{VAR}-1)^2} \quad (1)$$

$$\mathrm{BIAS} = \frac{\mu_e}{\mu_o} \quad (2)$$

$$CC = \frac{\sum_{i=1}^{n}(O_i - \bar{O})\,(E_i - \bar{E})}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}\cdot\sqrt{\sum_{i=1}^{n}(E_i - \bar{E})^2}} \quad (3)$$

$$VAR = \frac{CV_e}{CV_o} = \frac{\sigma_e/\mu_e}{\sigma_o/\mu_o} \quad (4)$$

where O is the value observed by the rain gauges, $\bar{O}$ is the mean gauged values, E is the value estimated by satellite-based products, $\bar{E}$ is the mean estimated values, μ is the distribution mean and σ is the standard deviation. The subscripts e and o correspond to the estimated and gauged data, respectively. KGE values range from -∞ to 1, with desirable values close to 1 and negative values representing worse performances.

A second evaluation step was to verify the performance of the rainfall threshold determined by each rainfall product to identify true or false alarms using a binary classifier of the rainfall conditions that do or do not lead to floods occurrences. The same contingency matrix applied by Ramos Filho et al. (2021), consisting of four components, was used for each threshold, including: 1) true positive (TP) when the threshold is exceeded and the flood occurs, 2) false negative (FN) when the threshold is not exceeded and the flood occurs, 3) false positive (FP) when the threshold is exceeded and the flood does not occur, and 4) true negative (TN) when the threshold is not exceeded and the flood does not occur. The following three metrics were then applied using the above-mentioned contingency matrix to assess the skill score of the floods thresholds: 1) probability of detection (POD), which measures the fraction of events that are correctly predicted by the satellite-based products; 2) false alarm ratio (FAR), which exhibits the fraction of events incorrectly detected by the satellite-based products; and 3) Hanssen-Kuiper (HK) skill score, which measures the applicability/quality to identify the usability and accuracy of the threshold:

$$POD = \frac{TP}{TP + FN} \quad (3)$$

$$FAR = \frac{FP}{FP + TN} \quad (4)$$

$$HK = POD - FAR \quad (5)$$

The values of these three metrics range from 0% to 100%. The perfect values for POD and HK are close to 100%, while the desirable values for FAR are close to 0%.

4. Results and discussion

4.1 Characterisation of rainfall events that trigger floods

Figure 2 presents the classifications of the daily rainfall considering the values registered by the rain gauges and the analysed products only for the days where floods events were registered. The results show that the ground-based data presented only heavy (45%) and severe (55%) rain records during the analysed period. Conversely, all satellite-based rainfall products presented light rain and moderate rain records ranging from 19% (CHIRPS V2.0) to 53% (SM2RAIN-ASCAT V1.2) and from 36% (MERRA-2) to 62% (GPM+SM2RAIN). This indicates an underestimation of the daily accumulated precipitation by all analysed products. Among the dataset able to detect daily heavy and severe rains when flood occurrences were registered, the following products stand out: 1) PDIR-Now, represented by 31% of heavy rain and 6% of severe rain; 2) CMORPH-CRT V1.0, characterised by 32% of heavy rain and 5% of severe rain; 3) PERSIANN-CCS, which presented 26% of heavy rain and 6% of severe rain; and 4) IMERG-F, showing 25% of heavy rain and 5% of severe rain.

INSERT FIG. 2 HERE

Fig. 2. Daily precipitation classification that leads to flood occurrences in São Paulo State. The symbols * and ** highlight the uncorrected (solely satellite data) and corrected (satellite and gauge data) products, respectively.

Satellite-based underestimations of these extreme precipitation events, when compared with the rain gauge observations, were also reported by other researchers previously for a variety of products (e.g., Mayor et al., 2017; Solakian et al., 2020; Xuan et al., 2020). Thus, it is important to analyse the performance of multiple precipitation products over the region of interest instead of relying on randomly chosen products (Masunaga et al., 2019), because the performance of these products in capturing the spatiotemporal variability of extremes rainfall depends on season, regions, time period, and inexistence or scarcity of rain gauges to bias-correct products (Chen et al., 2020).

Figure 3 presents the mean KGE scores of the 14 satellite-based rainfall products considering the six accumulated rainfall periods. The results of the KGE show that all products presented negative mean scores for time steps ranging from 3 h to 1 day (-0.64 to -0.41, on average). The best and worst performances of the mean KGE scores for the sub-daily dataset, considering these first four aggregation periods (from 3 h to 1 day), were identified to be the IMERG-F (i.e., -0.38) and MERRA-2 (i.e., -1.28) products, respectively. When only the daily datasets are considered, the best and worst performances of the daily mean KGE scores were observed as the GPM+SM2RAIN (i.e., -0.08) and PERSIANN-CDR V1R1 (i.e., -0.44) products, respectively. Overall, it is noticeable an improvement in the KGE values when longer rainfall accumulation times are considered. Null to positive mean KGE scores were observed for the time steps of 3 and 10 days (-0.02 to 0.18, on average). Variability was the main responsible for the poor performance for time steps varying between 3 hours and 1 day (2.12 to 1.87, on average) presenting values far from the ideal (1). Moreover, the BIAS (0.25 – 0.27, on average) and CC (0.08 to 0.19, on average) also presented their worst results for such time steps. For longer time steps (3 days and 10 days), the variability presented results closer to ideal (1.45 – 1.18, on average), while the BIAS (0.39 – 0.57, on average) and CC (0.34 – 0.35, on average) remained furthest from the desirable values for all products.

INSERT FIG. 3 HERE

Fig. 3. Graph showing the (a) KGE, (b) CC, (c) BIAS, and (d) VAR scores for the 14 satellite-based rainfall, considering only extreme precipitation events for time steps ranging from 3 hours to 10 days. The red lines represent the perfect values. The symbols * and ** highlight the uncorrected (solely satellite data) and corrected (satellite and gauge data) products, respectively.

Overall, the best values of the analysed metrics were found for all products when the rainfall was accumulated over 10 days. Two gauge-based uncorrected products (PDIR-Now and GPM+SM2RAIN) presented the highest values of KGE (i.e., 0.36) for the time step equal to 10 days. However, PDIR-Now presented the highest BIAS (i.e., 076), indicating that this product better represents the total precipitation compared to GPM+SM2RAIN (BIAS = 0.54). On the other hand, the data from GPM+SM2RAIN better linearly correlates with the gauged data (CC = 0.55) when compared to the PDIR-Now product. The performance of these two above-mentioned uncorrected products was followed by the following corrected products: CHIRPS V2.0 (KGE = 0.34), MSWEP V2.2 (KGE = 0.28), and IMERG-F (KGE = 0.27).

4.1.1. Overall analysis of the uncorrected dataset

Among the eight uncorrected satellite products (i.e., PERSIANN, PERSIANN-CCS, PDIR-Now, SM2RAIN-ASCAT, GPM+SM2RAIN, IMERG-E, IMERG-L, and CHIRP V2.0), the GPM+SM2RAIN product performed better for extreme precipitations when considering the daily time step onwards, with a mean KGE value of 0.15, followed by PDIR-Now and CHIRP V2.0, with KGE values equal to 0.03 for both. PDIR-Now is a product, intended to replace the PERSIANN-CCS datasets, which considers the errors and uncertainties resulting from the use of infrared images (Nguyen et al., 2020). Nevertheless, PERSIANN-CCS performed slightly better than PDIR-Now for sub-daily time steps, reversing the position for daily time steps onward. The PERSIANN product presented the lowest values of KGE among the uncorrected analysed products, ranging from -0.60 (3 hours) to 0.04 (10 days). All sub-daily uncorrected products presented extremely low values of KGE for time steps below 1 day, with means ranging from -0.76 (PERSIANN) to -0.45 (PERSIANN-CCS). Overall, the GPM+SM2RAIN product performed noticeably better than SM2RAIN-ASCAT V1.2 (mean KGE = -0.11), i.e., the other product that also uses satellite-based soil moisture data. The two microwave-based datasets (IMERG-E and IMERG-L) showed similar results for all analysed time steps, with mean KGE values equal to -0.05 when considering daily onward time steps, i.e., slightly worse than that observed for CHIRP V2.0 (KGE = 0.03).

4.1.2. Overall analysis of the corrected dataset

The products corrected by ground observations use daily, 5-day, 10-day, and/or monthly precipitation data in their algorithms. Among the gauged corrected products, CHIRPS V2.0 presented the higher values of KGE for extreme precipitations, varying for daily time step onwards between -0.28 (1-day) and 0.34 (10-days). The performance of this product was followed by MSWEP V2.2 and IMERG-F, with KGE values ranging from -0.65 (3-hours) to 0.28 (10-days) and from -0.45 (3-hours) to 0.27 (10-days), respectively. The CMORPH-CRT V1.0 product presented a similar performance to those observed for MSWEP V2.2 and IMERG-F, with KGE values varying between -0.47 (3-hours) and 0.26 (10-days). MERRA-2 and CMORPH-CRT V1.0 exhibited the lowest values of KGE among the products corrected by ground observations, with overall performances even worse than all those observed for the uncorrected products. The performance of these products may be affected by some factors in rain gauges, such as miscellaneous technical errors, different reporting times, different quality control procedures, network density, among others (Beck et al., 2019a; Derin and Yilmaz, 2014; Shen et al., 2021; Sun et al., 2018).

4.2 Rainfall thresholds

4.2.1 Evaluation for different tolerance levels

Figure 4 shows the heatmaps with the main values of POD, FAR, and HK for the six considered tolerance levels. The values of POD are set by the adopted tolerance levels, with values varying between 0.95 and 0.50 for the no-exceedances probabilities of 5 and 50% for all satellite-based and the gauged data, respectively (Fig. 4a). On the other hand, the FAR values presented reductions as the tolerance levels increased, with the worst performance observed for CHIRPS V2.0, PERSIANN and PERSIANN-CDR

V1R1 (FAR≈0.75) adopting a tolerance level of 5% tolerance (Fig. 4b). The rainfall products CMORPH-CRT V1.0, IMERG-F, and GPM+SM2RAIN exhibited the lower values of FAR (0.11) when a tolerance level of 50% was considered, i.e., like the gauged data but with a tolerance level of 5% only (FAR = 0.13). The product that performed better overall in the number of false alarms (IMERG-F) showed values varying between 0.50 (5%) and 0.11 (50%), i.e., much higher than those observed for the rain gauges, which ranged from 0.13 to 0.02, respectively.

INSERT FIG. 4 HERE

Fig. 4. Heatmap of the mean values of (a) POD, (b) FAR, and (c) HK using different no-exceedance probability. The symbols * and ** highlight the uncorrected (solely satellite data) and corrected (satellite and gauge data) products, respectively.

Overall, all analysed products showed similar patterns in HK, with an increase in the values until a certain peak value, mostly between the application of tolerance levels of 10% and 30%, before a decline in the values of this metric for higher tolerance levels. The difference is that the CHIRPS V2.0, PERSIANN, PERSIAN-CCS, and PERSIAN-CDR V1R1 products exhibited peak values of HK for tolerance levels varying between 30 and 40%. The HK values indicate better performance for the gauged data utilising a tolerance level of 5% (HK = 0.83), followed by the CHIRP V2.0 and IMERG-F products, which presented HK values equal to 0.51 and 0.52, respectively, for tolerance levels of 20%. Although presenting the highest values of HK, these two mentioned products still exhibited a considerable rate of false alarms, around 28%. Moreover, CHIRPS V2.0 and PERSIANN-CDR V1R1 also had the worst performance for this metric, as expected, with the highest HK values equal to 0.29 and 0.36 for a tolerance level of 30%, respectively. The study carried out by Brunetti et al. (2018), which analysed 4 satellite products for delimitation of landslide thresholds in Italy, showed that the SM2RAIN-ASCAT V1.2 product presented the highest values of HK equals to 0.42 for exceedance limits between 20-25%, while the PERSIANN product performed worse, with HK value equals to 0.31 for a tolerance level of 25%. Jia et al. (2020) also analysed 4 rainfall products for landslide thresholds on a global scale, including CMORPH, which better performed among the evaluated, presenting an HK value equal to 0.43 for a tolerance level of 22%. The same global scale study identified that PERSIANN presented the lowest values of HK among the analysed products, with the best result (HK = 0.14) found for a tolerance level equal to 9%.

4.2.2 Determination of rainfall thresholds

Figure 5 presents the six precipitation thresholds obtained from the tolerance limits of 5, 10, 20, 30, 40, and 50% for the gauged data and all analysed precipitation products. It is possible to observe a considerable underestimation of the thresholds delineated by the satellite-based products compared to those elaborated by the gauged data, with larger differences noticed for shorter time steps (3h–1d), which were smoothed for longer considered periods (3–10 days). For instance, the CHIRPS V2.0, GPM+SM2RAIN, and PDIR products presented, respectively, values for 10-days accumulated rainfalls equal to 54.6, 53.0, and 67.0 mm considering a tolerance level of 20%, which correspond, respectively, to biases of 0.6, 0.58, and 0.72 (Figure 6) when compared to the thresholds using the gauged data (i.e., 91,8 mm). This behaviour was expected as it is difficult for the satellite-based rainfall products to capture the precipitation peaks, resulting in values with more space in time compared to those identified for the gauged dataset. The worst performances of the satellite-based products for longer time steps (i.e., daily onwards) were verified for 3 days accumulated rainfall considering tolerance levels of 5%, with PERSIANN (2.71 mm) and MSWEP V2.2 (3.73 mm) presenting the largest differences to the gauged data (42.6 mm), i.e., biases of 0.06 and 0.09, respectively. For shorter time-steps ranging from 3h to 1d, the worst results were found for tolerance levels of 5%, where almost all products presented BIAS values equal to or lower than 0.1. An exception, considering a tolerance level of 5%, was observed for the CHIRPS products, with a BIAS value equal to 0.17 for accumulated rainfall of 1 day (6.41 mm, i.e., still much lower than the 37 mm obtained with the gauged data). The best results found for shorter time-steps were verified for 1-day accumulated rainfall considering tolerance levels of 50% tolerance level. For this combination of accumulated rainfall and tolerance level, for instance, the products CMORPH-CRT V1.0, IMERG-F V06, and PDIR presented values equal to 22.75, 20.17, and 20.0 mm, while the gauged data, for the same time-step and non-exceedance probability, exhibited a value of 73.5 mm. Overall, it is noticed in Figure 6 that the time steps presented greater relevance in the improvement of the BIAS values compared to the tolerance limits.

INSERT FIG. 5 HERE

Fig. 5. Accumulated precipitation versus duration applying the tolerance levels of 5, 10, 20, 30, 40, and 50% for the (a) rain gauges, (b) CHIRP V2.0, (c) CHIRPS V2.0, (d) IMERG-E V06, (e) IMERG-L V06, (f) IMERG-F V06, (g) CMORPH-CRT V01, (h) MERRA-2, (i) MSWEP V2.2, (j) PERSIANN, (k) PERSIANN-CCS, (l) PERSIANN-CDR V1R1, (m) PDIR-Now, (n) SM2RAIN-ASCAT V1.2, and (o)

GPM+SM2RAIN. The symbols * and ** highlight the uncorrected (solely satellite data) and corrected (satellite and gauge data) products, respectively.

INSERT FIG. 6 HERE

Fig. 6. BIAS values for the estimated rainfall thresholds, using rain gauge as a reference, for tolerance levels of (a) 5, (b) 10, (c) 20, (d) 30, (e) 40, and (f) 50%. The red lines represent the perfect values. The symbols * and ** highlight the uncorrected (solely satellite data) and corrected (satellite and gauge data) products, respectively.

5. Conclusions

In this study, sub(daily) rainfall data from 14 different satellite-based products were evaluated to characterise rainfall events that trigger floods. A collection and filtering of observed information from 583 rain gauges and 551 flood occurrences were also used for this evaluation. The gauged values were compared with the precipitation products for 6 different time steps (3 h, 6 h, 12 h, 1 d, 3 d, and 10 d). The applicability of a methodology for determining precipitation thresholds using satellite-based products was also evaluated. The two main findings of this study are summarised as follows:

(1) Overall, all analysed products tend to largely underestimate the extreme rainfall events (i.e., when and where flood events were registered) observed by the rain gauges, mainly at sub-daily scales. This underestimation primarily occurred due to the difficulty of the estimated products to capture precipitation peaks, as their values are more distributed over time with longer durations. The point-to-pixel analysis used in this study tended to contribute more to the underestimation of the gauged peak intensity due to the representation of a spatial average of precipitation at the pixel scale. The best results evaluating the extreme rainfall events were expected for products corrected by ground-based rainfall stations, but they were found for the PDIR-Now and GPM+SM2RAIN products considering 10-days accumulated precipitation, followed by the corrected CHIRPS V2.0, MSWEP V2.2, and GPM IMERG-F, although the results (i.e., KGE values ranging from 0.36 to 0.27) indicate that all products are far from ideal (KGE=1).

(2) Large underestimations were also identified for the rainfall thresholds delineated by the satellite-based products. Despite the large underestimations, the delineation of rainfall thresholds using satellite-based products is possible but with lower precipitation values and a greater probability of false alarm occurrences. The gauged rainfall data, considering tolerance limits of 5%, presented mean HK and

BIAS values for daily rainfall data ~60 and 65%, respectively, higher than the two products that better delineated the rainfall thresholds (e.g., CHIRP V2.0 and IMERG-F) but considering tolerance levels of 20%.

Based on our findings, the rainfall satellite-based products dataset, even less accurate than the ground-based observations, can be applied, when multi-daily accumulated data are considered, as an alternative source of data for determining precipitation thresholds in some regions that present low-density of rain gauges. PDIR-Now showed to be an interesting source of data to characterise flood events since this product provides near-real-time information, followed by the SM2RAIN products, which correct the satellite rainfall data without the use of ground-based information. For regions with a high density of rain gauges with sub-daily data available, the use of ground-based data will still provide a much better source of information to characterise events that trigger floods. Therefore, the use of new approaches (e.g., merging of products or improvement of algorithms) must be explored to enable better identification and characterisation of extreme rainfall events over areas with low availability of in-situ sub-daily data and, consequently, improve the delineation of thresholds for monitoring flood hazards. Also, some sources of uncertainties in the analysis of this study (e.g., the biases toward the rain gauges, as the event's selection is done by considering the rainfall amount and the distance from rain gauges) require a better assessment, which includes the performance of: (1) different analysis (i.e., a different set of events) when rain gauge and satellite data are used, and (2) bias correction of satellite data as they represent area-averaged rainfall not directly comparable with point measurements from the rain gauge.

Acknowledgements: We acknowledge CEMADEN (Brazilian Centre for Monitoring and Early Warnings of Natural Disasters), SIMPAT (Integrated Storm Monitoring, Forecasting and Alerting System for the Brazilian South-Southeast Regions), and the Civil Defence of the state of São Paulo for providing the rainfall and occurrences dataset. We also would like to thank the financial support granted by 1) the Research Support Foundation of Paraíba State (FAPESQ-PB), which funded the contribution from Yunqing Xuan, supported in partnership with the Newton Fund, via CONFAP – The UK Academies Research Mobility 2017/2018 (Grant REF: 039/2018); 2) the Brazilian National Council for Scientific and Technological Development (CNPq), which funds the Universal MCTI/CNPq No. 28/2018 (Grant REF: 433801/2018-2); and 3) the Brazilian Coordination for Improvement of Higher Education Personnel

(CAPES) – Finance Code 001. The authors also thanks the two anonymous reviewers and the editor for the constructive comments during the review process of this manuscript.

References

AghaKouchak, A., Farahmand, A., Melton, F.S., Teixeira, J., Anderson, M.C., Wardlow, B.D., Hain, C.R., 2015. Remote sensing of drought: Progress, challenges and opportunities. Rev. Geophys. 53, 452–480. https://doi.org/10.1002/2014RG000456

Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. Eng. Geol. 73, 247–265. https://doi.org/10.1016/j.enggeo.2004.01.007

Almagro, A., Oliveira, P.T.S., Brocca, L., 2021. Assessment of bottom-up satellite rainfall products on estimating river discharge and hydrologic signatures in Brazilian catchments. J. Hydrol. 603. https://doi.org/10.1016/j.jhydrol.2021.126897

Alvares, C.A., Stape, J.L., Sentelhas, P.C., De Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Zeitschrift 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Ashouri, H., Hsu, K.L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull. Am. Meteorol. Soc. 96, 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1

Beck, H.E., Pan, M., Roy, T., Weedon, G.P., Pappenberger, F., Van Dijk, A.I.J.M., Huffman, G.J., Adler, R.F., Wood, E.F., 2019a. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. Hydrol. Earth Syst. Sci. 23, 207–224. https://doi.org/10.5194/hess-23-207-2019

Beck, H.E., Van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., De Roo, A., 2017a. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21, 589–615. https://doi.org/10.5194/hess-21-589-2017

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A.I.J.M., Weedon, G.P., Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017b. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Adv. Glob. Chang. Res. 69, 625–653. https://doi.org/10.1007/978-3-030-35798-6_9

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2019b. MSWep v2 Global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 100, 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Berti, M., Martina, M.L.V., Franceschini, S., Pignone, S., Simoni, A., Pizziolo, M., 2012. Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. J. Geophys. Res. Earth Surf. 117, 1–20. https://doi.org/10.1029/2012JF002367

Blenkinsop, S., Fowler, H.J., Barbero, R., Chan, S.C., Guerreiro, S.B., Kendon, E., Lenderink, G., Lewis, E., Li, X.-F., Westra, S., Alexander, L., Allan, R.P., Berg, P., Dunn, R.J.H., Ekström, M., Evans, J.P., Holland, G., Jones, R., Kjellström, E., Klein-Tank, A., Lettenmaier, D., Mishra, V., Prein, A.F., Sheffield, J., Tye, M.R., 2018. The INTENSE project: using observations and models to understand the past, present and future of sub-daily rainfall extremes. Adv. Sci. Res. 15, 117–126. https://doi.org/10.5194/asr-15-117-2018

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., Wagner, W., 2019. SM2RAIN-ASCAT (2007-2018): Global daily satellite rainfall data from ASCAT soil moisture observations. Earth Syst. Sci. Data 11, 1583–1601. https://doi.org/10.5194/essd-11-1583-2019

Brocca, L., Massari, C., Tarpanelli, A., Brocca, L., Crow, W.T., Ciabatta, L., Massari, C., Rosnay, P. De, Enenkel, M., Hahn, S., Amarnath, G., Camici, S., Tarpanelli, A., Wagner, W., Member, S., De Rosnay, P., Enenkel, M., Hahn, S., Amarnath, G., Camici, S., Tarpanelli, A., Wagner, W., 2017. A Review of the Applications of ASCAT Soil Moisture Products A Review of the Applications of ASCAT Soil Moisture Products. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2285–2306.

Brollo, M.J., Ferreira, C.J., 2016. Gestão de risco de desastres devido a fenômenos geodinâmicos no estado de São Paulo: Cenário 2000-2015, Boletim do Instituto Geológico nº 67. Instituto Geológico, São Paulo, Brazil. https://doi.org/10.13140/RG.2.2.34788.17288

Brunetti, M.T., Melillo, M., Gariano, S.L., Ciabatta, L., Brocca, L., Amarnath, G., Peruccacci, S., 2021. Satellite rainfall products outperform ground observations for landslide prediction in India. Hydrol. Earth Syst. Sci. 25, 3267–3279. https://doi.org/10.5194/hess-25-3267-2021

Brunetti, M.T., Melillo, M., Peruccacci, S., Ciabatta, L., Brocca, L., 2018. How far are we from the use of satellite rainfall products in landslide forecasting? Remote Sens. Environ. 210, 65–75. https://doi.org/10.1016/j.rse.2018.03.016

Camici, S., Massari, C., Ciabatta, L., Marchesini, I., Brocca, L., 2020. Which rainfall score is more informative about the performance in river discharge simulation? A comprehensive assessment on 1318 basins over Europe. Hydrol. Earth Syst. Sci. 24, 4869–4885. https://doi.org/10.5194/hess-24-4869-2020

Chen, J., Li, Z., Li, L., Wang, J., Qi, W., Xu, C.Y., Kim, J.S., 2020. Evaluation of multi-satellite precipitation datasets and their error propagation in hydrological modeling in a monsoon-prone region. Remote Sens. 12, 1–33. https://doi.org/10.3390/rs12213550

Chikalamo, E.E., Mavrouli, O.C., Ettema, J., van Westen, C.J., Muntohar, A.S., Mustofa, A., 2020. Satellite-derived rainfall thresholds for landslide early warning in Bogowonto Catchment, Central Java, Indonesia. Int. J. Appl. Earth Obs. Geoinf. 89, 102093. https://doi.org/10.1016/j.jag.2020.102093

Chikoore, H., Bopape, M.-J.M., Ndarana, T., Muofhe, T.P., Gijben, M., Munyai, R.B., Manyanya, T.C., Maisha, R., 2021. Synoptic structure of a sub-daily extreme precipitation and flood event in Thohoyandou, north-eastern South Africa. Weather Clim. Extrem. 33, 100327. https://doi.org/10.1016/j.wace.2021.100327

Derin, Y., Yilmaz, K.K., 2014. Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. J. Hydrometeorol. 15, 1498–1516. https://doi.org/10.1175/JHM-D-13-0191.1

Diakakis, M., 2012. Rainfall thresholds for flood triggering. The case of Marathonas in Greece. Nat. Hazards 60, 789–800. https://doi.org/10.1007/s11069-011-9904-7

Dinis, P.A., Huvi, J., Cabral Pinto, M., Carvalho, J., 2021. Disastrous Flash Floods Triggered by Moderate to Minor Rainfall Events. Recent Cases in Coastal Benguela (Angola). Hydrology 8, 73. https://doi.org/10.3390/hydrology8020073

Du, S., Shi, P., Van Rompaey, A., Wen, J., 2015. Quantifying the impact of impervious surface location on flood peak discharge in urban areas. Nat. Hazards 76, 1457–1471. https://doi.org/10.1007/s11069-014-1463-2

Dunkerley, D., 2019. Sub-daily rainfall intensity extremes: Evaluating suitable indices at Australian arid and wet tropical observing sites. Water (Switzerland) 11. https://doi.org/10.3390/w11122616

Freitas, E. da S., Coelho, V.H.R., Xuan, Y., Melo, D. de C.D., Gadelha, A.N., Santos, E.A., Galvão, C. de O., Ramos Filho, G.M., Barbosa, L.R., Huffman, G.J., Petersen, W.A., Almeida, C. das N., 2020. The performance of the IMERG satellite-based product in identifying sub-daily rainfall events and their properties. J. Hydrol. 589, 125128. https://doi.org/10.1016/j.jhydrol.2020.125128

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C., Martius, O., 2015. Flood triggering in Switzerland: The role of daily to monthly preceding precipitation. Hydrol. Earth Syst. Sci. 19, 3903–3924. https://doi.org/10.5194/hess-19-3903-2015

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. Sci. Data 2, 1–21. https://doi.org/10.1038/sdata.2015.66

Gadelha, A.N., Coelho, V.H.R., Xavier, A.C., Barbosa, L.R., Melo, D.C.D., Xuan, Y., Huffman, G.J., Petersen, W.A., Almeida, C.N., 2018. Grid box-level evaluation of IMERG over Brazil at various space and time scales. Atmos. Res. #pagerange#. https://doi.org/10.1016/J.ATMOSRES.2018.12.001

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Getirana, A., Kirschbaum, D., Mandarino, F., Ottoni, M., Khan, S., Arsenault, K., 2020. Potential of GPM IMERG precipitation estimates to monitor natural disaster triggers in urban areas: The case of Rio de Janeiro, Brazil. Remote Sens. 12, 1–20. https://doi.org/10.3390/rs12244095

Glade, T., Crozier, M., Smith, P., 2000. Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "Antecedent Daily Rainfall Model." Pure Appl. Geophys. 157, 1059–1079. https://doi.org/10.1007/s000240050017

Gupta, H. V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hallegatte, S., Green, C., Nicholls, R.J., Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. Nat. Clim. Chang. 3, 802–806. https://doi.org/10.1038/nclimate1979

He, S., Wang, J., Liu, S., 2020. Rainfall event-duration thresholds for landslide occurrences in China. Water (Switzerland) 12. https://doi.org/10.3390/w12020494

Hegerl, G.C., Black, E., Allan, R.P., Ingram, W.J., Polson, D., Trenberth, K.E., Chadwick, R.S., Arkin, P.A., Sarojini, B.B., Becker, A., Dai, A., Durack, P.J., Easterling, D., Fowler, H.J., Kendon, E.J., Huffman, G.J., Liu, C., Marsh, R., New, M., Osborn, T.J., Skliris, N., Stott, P.A., Vidale, P.L., Wijffels, S.E., Wilcox, L.J., Willett, K.M., Zhang, X., 2015. Challenges in quantifying changes in the global water cycle. Bull. Am. Meteorol. Soc. 96, 1097–1115. https://doi.org/10.1175/BAMS-D-13-00212.1

Hong, Y., Hsu, K.L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteorol. 43, 1834–1852. https://doi.org/10.1175/jam2173.1

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.-L., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie, P., 2019. Algorithm Theoretical Basis Document (ATBD) Version 06 NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Natl. Aeronaut. Sp. Adm. 1–34.

IBGE - Instituto Brasileiro de Geografia e Estatística, 2021. (IBGE). [WWW Document]. URL https://cidades.ibge.gov.br/brasil/sp/ (accessed 7.16.19).

Jang, J.H., 2015. An advanced method to apply multiple rainfall thresholds for urban flood warnings. Water (Switzerland) 7, 6056–6078. https://doi.org/10.3390/w7116056

Jia, G., Tang, Q., Xu, X., 2020. Evaluating the performances of satellite-based rainfall data for global rainfall-induced landslide warnings. Landslides 17, 283–299. https://doi.org/10.1007/s10346-019-01277-6

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 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. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kha, D.D., Nhu, N.Y., Long, V.V., Van, D.T.H., 2020. Utility of GSMap precipitation and point scale in gauge measurements for stream flow modelling - A case study in lam river basin, Vietnam. J. Ecol. Eng. 21, 39–45. https://doi.org/10.12911/22998993/116350

Kidd, C., Becker, A., Huffman, G.J., Muller, C.L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D.B., 2017. So, how much of the Earth's surface is covered by rain gauges? Bull. Am. Meteorol. Soc. 98, 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Levizzani, V., Kidd, C., Aonashi, K., Bennartz, R., Ferraro, R.R., Huffman, G.J., Roca, R., Turk, F.J., Wang, N.Y., 2018. The activities of the international precipitation working group. Q. J. R. Meteorol. Soc. 144, 3–15. https://doi.org/10.1002/qj.3214

Lewis, E., Fowler, H., Alexander, L., Dunn, R., Mcclean, F., Barbero, R., Guerreiro, S., Li, X.F., Blenkinsop, S., 2019. GSDR: A global sub-daily rainfall dataset. J. Clim. 32, 4715–4729. https://doi.org/10.1175/JCLI-D-18-0143.1

Li, Z., Liu, H., 2020. Temporal and spatial variations of precipitation change from southeast to northwest china during the period 1961-2017. Water (Switzerland) 12. https://doi.org/10.3390/W12092622

Llauca, H., Lavado‐ casimiro, W., León, K., Jimenez, J., Traverso, K., Rau, P., 2021. Assessing near real-time satellite precipitation products for flood simulations at sub‐ daily scales in a sparsely gauged watershed in Peruvian andes. Remote Sens. 13, 1–18. https://doi.org/10.3390/rs13040826

Massari, C., Brocca, L., Pellarin, T., Abramowitz, G., Filippucci, P., Ciabatta, L., Maggioni, V., Kerr, Y., Fernandez Prieto, D., 2020. A daily 25km short-latency rainfall product for data-scarce regions based on the integration of the Global Precipitation Measurement mission rainfall and multiple-satellite soil moisture products. Hydrol. Earth Syst. Sci. 24, 2687–2710. https://doi.org/10.5194/hess-24-2687-2020

Masunaga, H., Schröder, M., Furuzawa, F.A., Kummerow, C., Rustemeier, E., Schneider, U., 2019. Inter-product biases in global precipitation extremes. Environ. Res. Lett. 14. https://doi.org/10.1088/1748-9326/ab5da9

Mayor, Y.G., Tereshchenko, I., Fonseca-Hernández, M., Pantoja, D.A., Montes, J.M., 2017. Evaluation of error in IMERG precipitation estimates under different topographic conditions and temporal scales over Mexico. Remote Sens. 9, 1–18. https://doi.org/10.3390/rs9050503

Mirus, B.B., Becker, R.E., Baum, R.L., Smith, J.B., 2018. Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning. Landslides 15, 1909–1919. https://doi.org/10.1007/s10346-018-0995-z

Monsieurs, E., Dewitte, O., Demoulin, A., 2019. A susceptibility-based rainfall threshold approach for landslide occurrence. Nat. Hazards Earth Syst. Sci. 19, 775–789. https://doi.org/10.5194/nhess-19-775-2019

Nanda Pratama, G., Suwarman, R., Dewa Gede Agung Junnaedhi, I., Riawan, E., Anugrah, A., 2017. Comparison landslide-triggering rainfall threshold using satellite data: TRMM and GPM in South Bandung area. IOP Conf. Ser. Earth Environ. Sci. 71, 0–9. https://doi.org/10.1088/1755-1315/71/1/012003

Nguyen, P., Shearer, E.J., Ombadi, M., Gorooh, V.A., Hsu, K., Sorooshian, S., Logan, W.S., Ralph, M., 2020. PERSIANN dynamic infrared-rain rate model (PDIR) for high-resolution, real-time satellite precipitation estimation. Bull. Am. Meteorol. Soc. 101, E286–E302. https://doi.org/10.1175/BAMS-D-19-0118.1

Pandey, V., Srivastava, P.K., 2019. Integration of microwave and optical/infrared derived datasets for a drought hazard inventory in a sub-tropical region of India. Remote Sens. 11. https://doi.org/10.3390/rs11040439

Papagiannaki, K., Lagouvardos, K., Kotroni, V., Bezes, A., 2015. Flash flood occurrence and relation to the rainfall hazard in a highly urbanized area. Nat. Hazards Earth Syst. Sci. 15, 1859–1871. https://doi.org/10.5194/nhess-15-1859-2015

Parker, A.L., Castellazzi, P., Fuhrmann, T., Garthwaite, M.C., Featherstone, W.E., 2021. Article applications of satellite radar imagery for hazard monitoring: Insights from Australia. Remote Sens. 13, 1–25. https://doi.org/10.3390/rs13081422

Ramos Filho, G.M., Coelho, V.H.R., Freitas, E. da S., Xuan, Y., Almeida, C. das N., 2021. An improved rainfall-threshold approach for robust prediction and warning of flood and flash flood hazards. Nat. Hazards 105, 2409–2429. https://doi.org/10.1007/s11069-020-04405-x

Ranghetti, L., Cardarelli, E., Boschetti, M., Busetto, L., Fasola, M., 2018. Assessment of water management changes in the Italian rice paddies from 2000 to 2016 using satellite data: A contribution to agro-ecological studies. Remote Sens. 10, 80–90. https://doi.org/10.3390/rs10030416

Sampson, C.C., Smith, A.M., Bates, P.B., Neal, J.C., Alfieri, L., Freer, J.E., 2015. A high-resolution global flood hazard model. Water Resour. Res. 51, 7358–7381. https://doi.org/10.1002/2015WR016954

Santos, M., Fragoso, M., 2016. Precipitation thresholds for triggering floods in the Corgo basin, Portugal. Water (Switzerland) 8. https://doi.org/10.3390/w8090376

Scheevel, C.R., Baum, R.L., Mirus, B.B., Smith, J.B., 2017. Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington: U.S. Geological Survey Open-File Report 2017–1039. https://doi.org/10.3133/ofr20171039

Sheffield, J., Wood, E.F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., Verbist, K., 2018. Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions. Water Resour. Res. 54, 9724–9758. https://doi.org/10.1029/2017WR022437

Shen, Z., Yong, B., Gourley, J.J., Qi, W., 2021. Real-time bias adjustment for satellite-based precipitation estimates over Mainland China. J. Hydrol. 596, 126133. https://doi.org/10.1016/j.jhydrol.2021.126133

Shrestha, P.K., Shrestha, S., Ninsawat, S., 2019. How significant is sub-daily variability of rainfall for hydrological modelling of floods? A satellite based approach to sub-daily downscaling of gauged rainfall. Meteorol. Appl. 26, 288–299. https://doi.org/10.1002/met.1762

Singh, L., Saravanan, S., 2020. Satellite-derived GRACE groundwater storage variation in complex aquifer system in India. Sustain. Water Resour. Manag. 6. https://doi.org/10.1007/s40899-020-00399-3

Solakian, J., Maggioni, V., Godrej, A.N., 2020. On the Performance of Satellite-Based Precipitation Products in Simulating Streamflow and Water Quality During Hydrometeorological Extremes. Front. Environ. Sci. 8, 1–20. https://doi.org/10.3389/fenvs.2020.585451

Sorooshian, S., Hsu, K.L., Gao, X., Gupta, H. V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. Bull. Am. Meteorol. Soc. 81, 2035–2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2

Špitalar, M., Gourley, J.J., Lutoff, C., Kirstetter, P.E., Brilly, M., Carr, N., 2014. Analysis of flash flood parameters and human impacts in the US from 2006 to 2012. J. Hydrol. 519, 863–870. https://doi.org/10.1016/j.jhydrol.2014.07.004

Su, J., Lü, H., Zhu, Y., Cui, Y., Wang, X., 2019. Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China. Atmos. Res. 225, 17–29. https://doi.org/10.1016/j.atmosres.2019.03.025

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.L., 2018. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. Rev. Geophys. 56, 79–107. https://doi.org/10.1002/2017RG000574

Sungmin, O., Kirstetter, P.E., 2018. Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. Q. J. R. Meteorol. Soc. 144, 270–281. https://doi.org/10.1002/qj.3218

Tan, M.L., Duan, Z., 2017. Assessment of GPM and TRMM precipitation products over Singapore. Remote Sens. 9, 1–16. https://doi.org/10.3390/rs9070720

Tan, M.L., Latif, A.B., Pohl, C., Duan, Z., 2014. Streamflow modelling by remote sensing: A contribution to digital Earth. IOP Conf. Ser. Earth Environ. Sci. 18, 0–6. https://doi.org/10.1088/1755-1315/18/1/012060

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S., 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environ. Model. Softw. 90, 201–216. https://doi.org/10.1016/j.envsoft.2017.01.006

Tominaga, L.K., Santoro, J., Amaral, R. do., 2015. Desastres naturais: conhecer para prevenir, 3rd ed, Instituto Geológico. Instituto Geológico, São Paulo, Brazil. https://doi.org/978-8587235091

Tsakiris, G., 2014. Flood risk assessment: Concepts, modelling, applications. Nat. Hazards Earth Syst. Sci. 14, 1361–1369. https://doi.org/10.5194/nhess-14-1361-2014

United Nations Office for Disaster Risk Reduction, Centre for Research on the Epidemiology of Disasters, 2020. Human Cost of Disasters, Human Cost of Disasters. https://doi.org/10.18356/79b92774-en

Vasco, D.W., Farr, T.G., Jeanne, P., Doughty, C., Nico, P., 2019. Satellite-based monitoring of groundwater depletion in California's Central Valley. Sci. Rep. 9, 1–14. https://doi.org/10.1038/s41598-019-52371-7

Wang, C., Tang, G., Han, Z., Guo, X., Hong, Y., 2018. Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements. J. Hydrol. 564, 342–356. https://doi.org/10.1016/j.jhydrol.2018.06.064

Xie, P., Joyce, R., Wu, S., Yoo, S.H., Yarosh, Y., Sun, F., Lin, R., 2017. Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. J. Hydrometeorol. 18, 1617–1641. https://doi.org/10.1175/JHM-D-16-0168.1

Xuan, D., Hu, Q., Wang, Y., Yang, H., Li, L., Wang, L., 2020. Precipitation characteristic analysis of the Zhoushan Archipelago: From the view of MSWEP and rainfall merging. Water (Switzerland) 12. https://doi.org/10.3390/w12030829

Yang, T.H., Hwang, G. Do, Tsai, C.C., Ho, J.Y., 2016. Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts. Hydrol. Earth Syst. Sci. 20, 4731–4745. https://doi.org/10.5194/hess-20-4731-2016

Young, A., Bhattacharya, B., Zevenbergen, C., 2021. A rainfall threshold-based approach to early warnings in urban data-scarce regions: A case study of pluvial flooding in Alexandria, Egypt. J. Flood Risk Manag. 14, 1–16. https://doi.org/10.1111/jfr3.12702

Yuan, F., Zhang, L., Soe, K.M.W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., Liu, Y., 2019. Applications of TRMM- and GPM-era multiple- satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in Myanmar. Remote Sens. 11. https://doi.org/10.3390/rs11020140

Figure 1

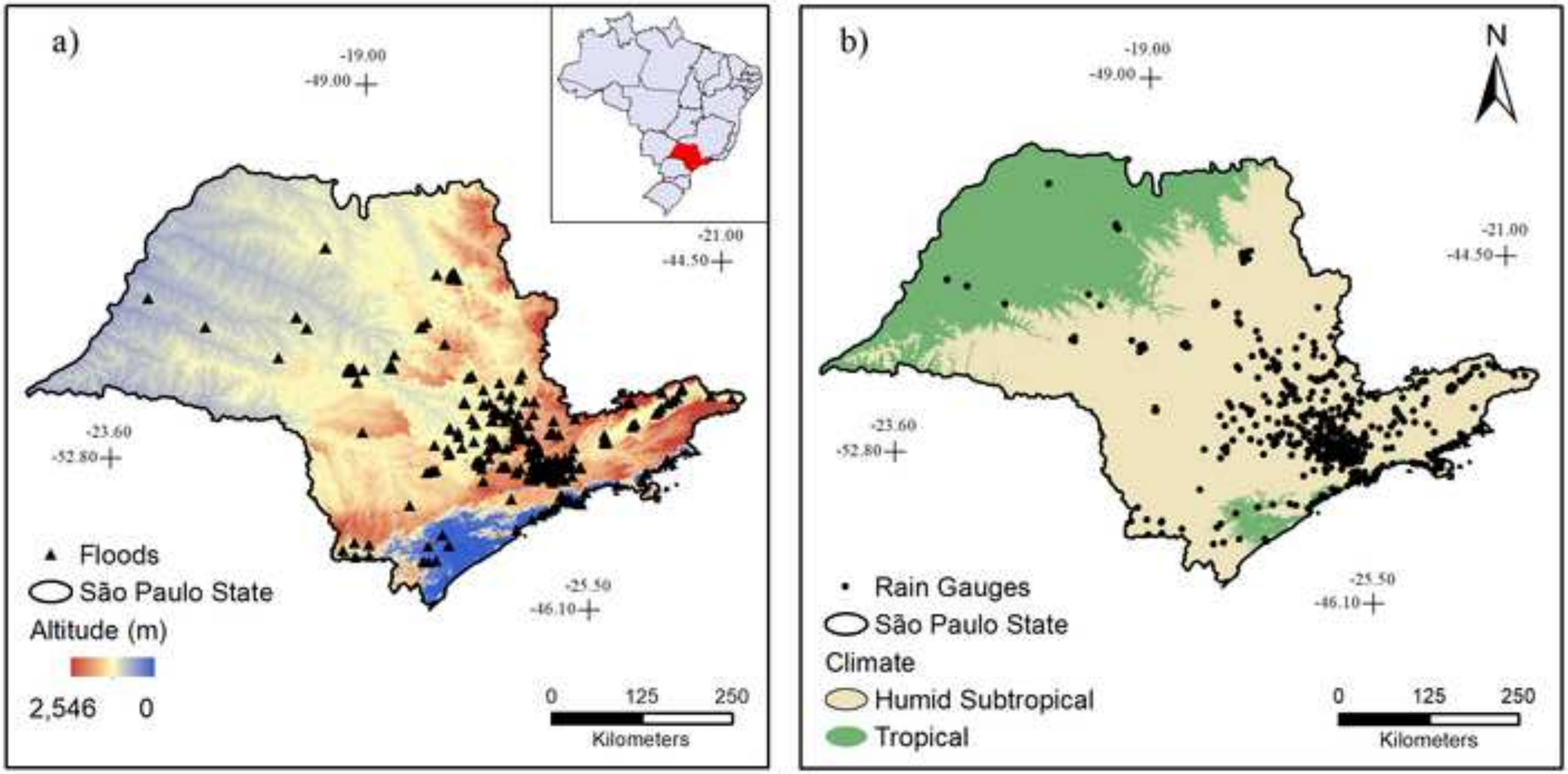
a)
-19.00
-49.00
-21.00
-44.50
-23.60
-52.80
-25.50
-46.10
Floods
São Paulo State
Altitude (m)
2,546
0
0
125
250
Kilometers
b)
N
-19.00
-49.00
-21.00
-44.50
-23.60
-52.80
-25.50
-46.10
Rain Gauges
São Paulo State
Climate
Humid Subtropical
Tropical
0
125
250
Kilometers

Figure 2

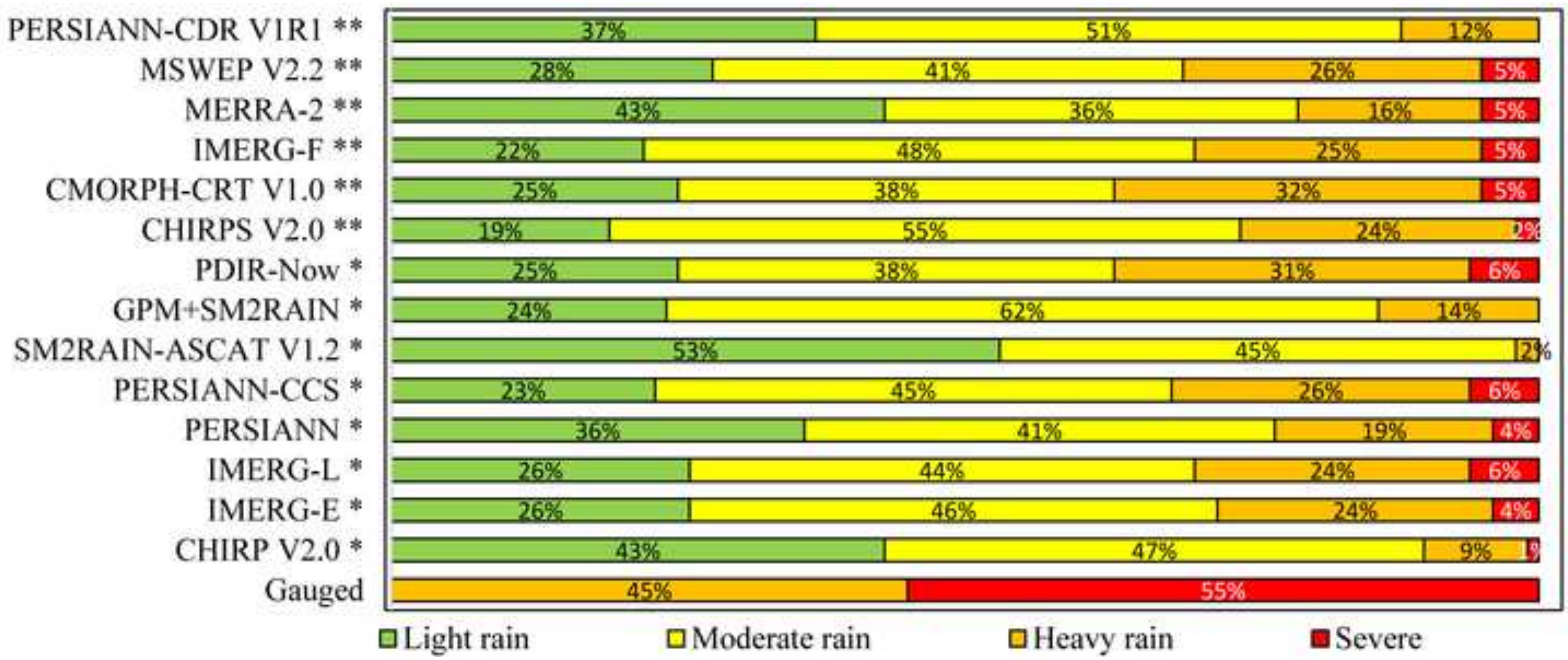
PERSIANN-CDR V1R1 **
MSWEP V2.2 **
MERRA-2 **
IMERG-F **
CMORPH-CRT V1.0 **
CHIRPS V2.0 **
PDIR-Now *
GPM+SM2RAIN *
SM2RAIN-ASCAT V1.2 *
PERSIANN-CCS *
PERSIANN *
IMERG-L *
IMERG-E *
CHIRP V2.0 *
Gauged
37% 51% 12%
28% 41% 26% 5%
43% 36% 16% 5%
22% 48% 25% 5%
25% 38% 32% 5%
19% 55% 24% 2%
25% 38% 31% 6%
24% 62% 14%
53% 45% 2%
23% 45% 26% 6%
36% 41% 19% 4%
26% 44% 24% 6%
26% 46% 24% 4%
43% 47% 9% 1%
45% 55%
Light rain
Moderate rain
Heavy rain
Severe

Figure 3

Click here to access/download;Figure;Fig3.tiff

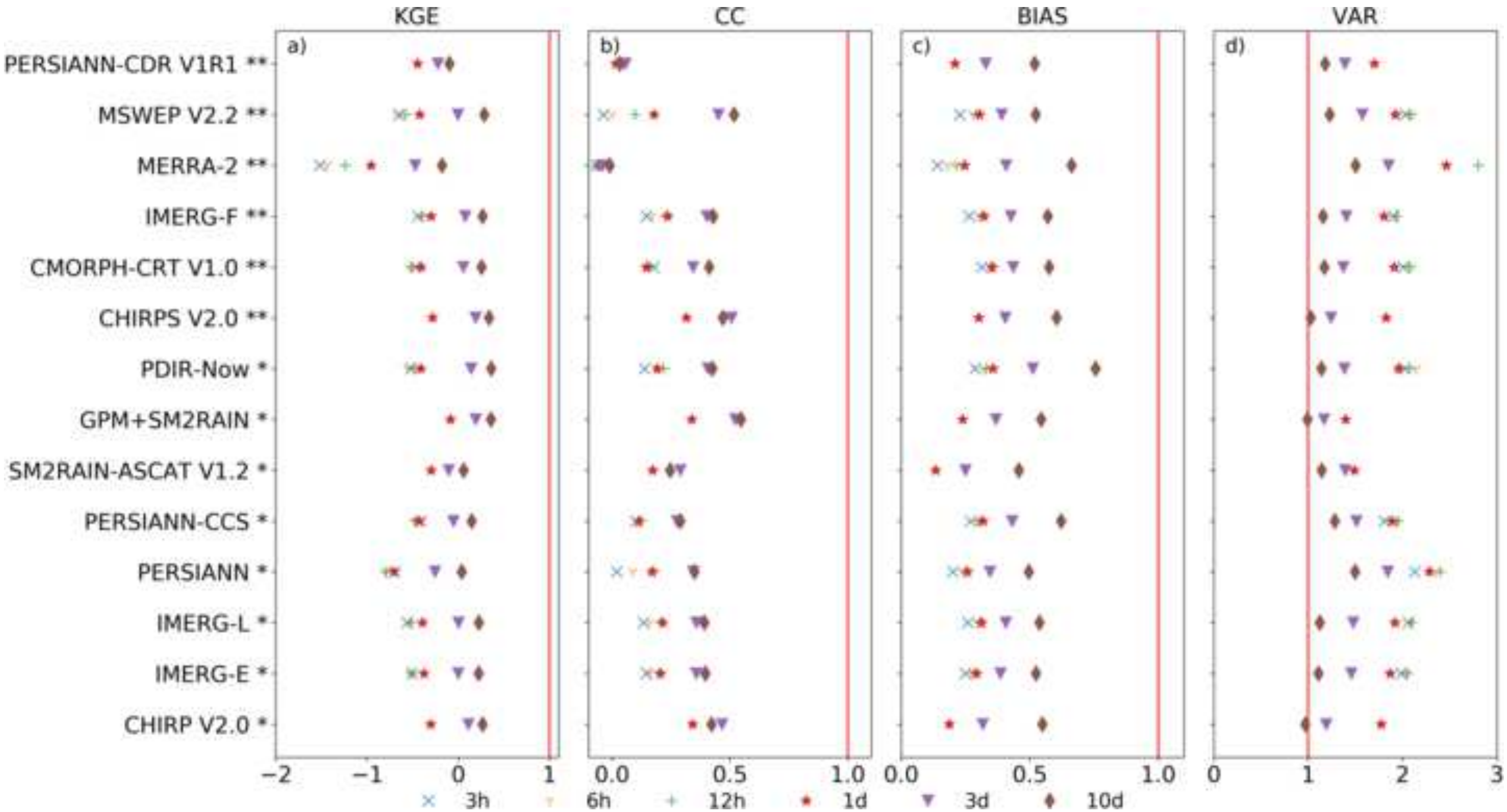

Figure 4

Click here to access/download;Figure;Fig4.tiff

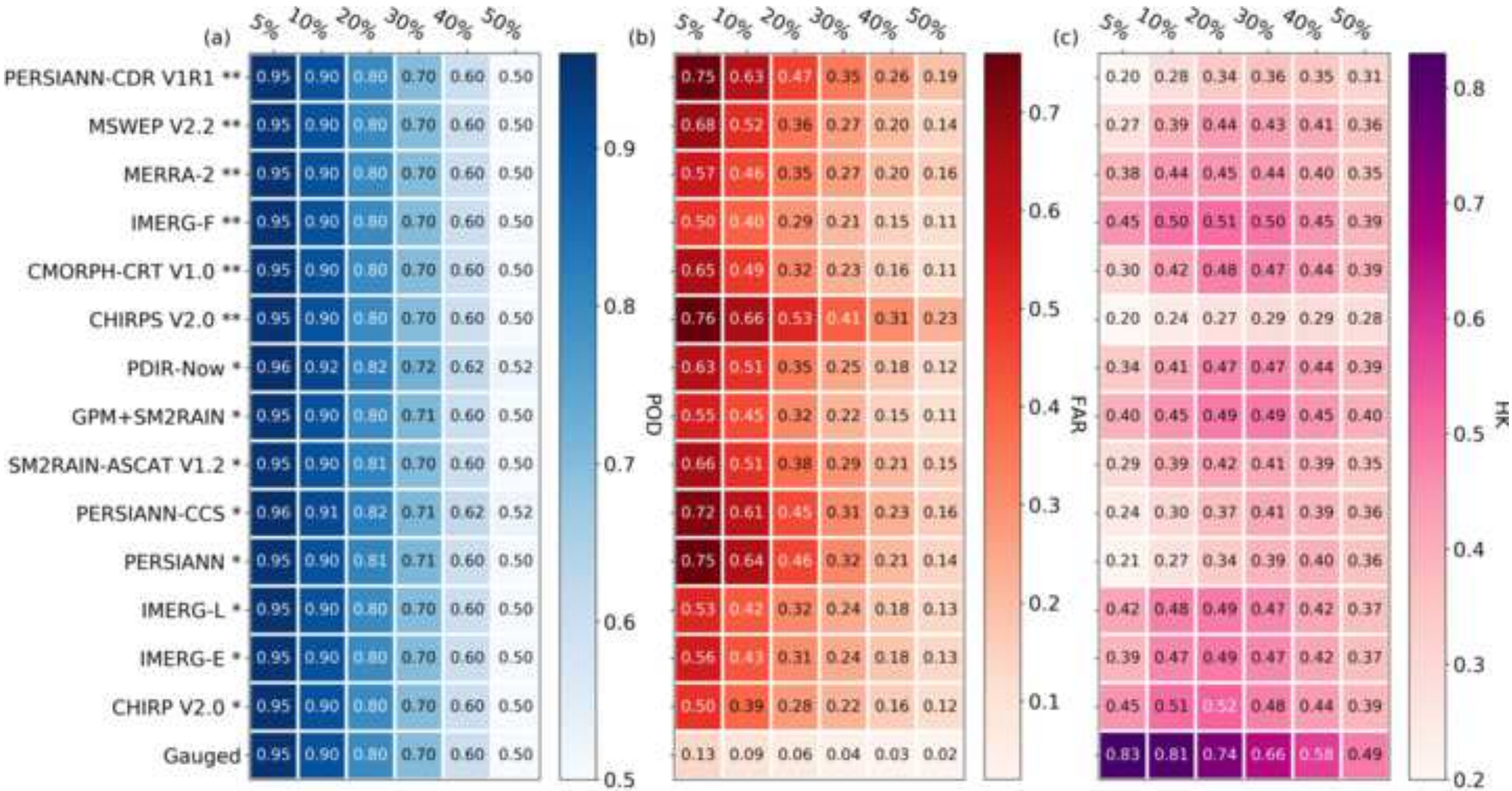

Figure 5 Click here to access/download;Figure;Fig5.tiff

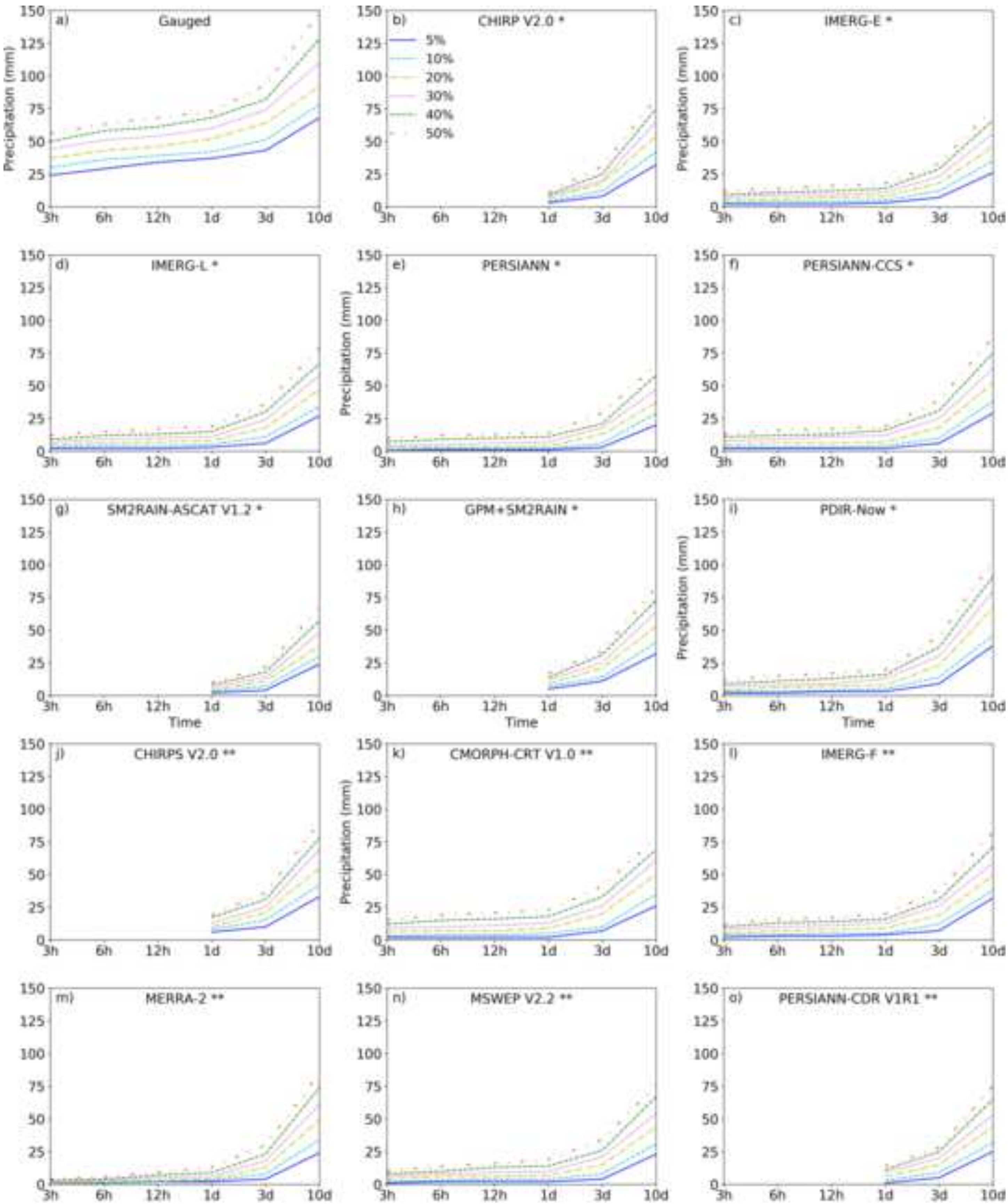

Figure 6

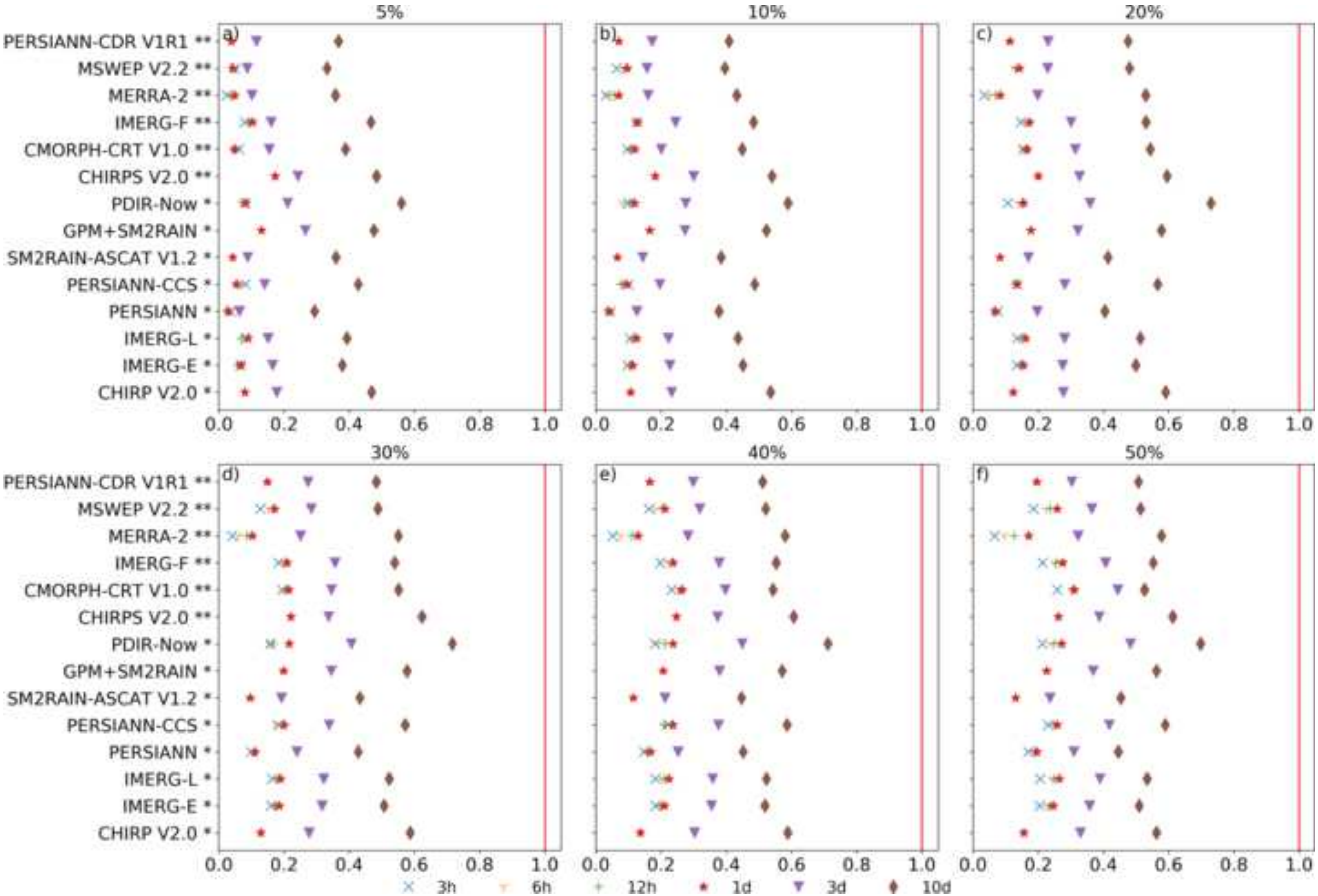
5%
10%
20%
30%
40%
50%
a)
b)
c)
d)
e)
f)
PERSIANN-CDR V1R1 **
MSWEP V2.2 **
MERRA-2 **
IMERG-F **
CMORPH-CRT V1.0 **
CHIRPS V2.0 **
PDIR-Now *
GPM+SM2RAIN *
SM2RAIN-ASCAT V1.2 *
PERSIANN-CCS *
PERSIANN *
IMERG-L *
IMERG-E *
CHIRP V2.0 *
0.0
0.2
0.4
0.6
0.8
1.0
3h
6h
12h
1d
3d
10d

Table 1

Name	Description	Spatial resolution	Spatial Coverage	Temporal Resolution	Temporal Coverage	Reference
CHIRP V2.0 *	Climate Hazards group InfraRed Precipitation (CHIRP) V2.0	0.05º	50º/S	Daily	1981-present	Funk et al. (2015)
IMERG-E V06 *	Integrated Multi-satellitE Retrievals for GPM (IMERG) early run V06	0.1º	60º N/S	30 min	2000-present	Huffman et al. (2019)
IMERG-L V06 *	Integrated Multi-satellitE Retrievals for GPM (IMERG) late run V06	0.1º	60º N/S	30 min	2000-present	Huffman et al. (2019)
PERSIANN *	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN)	0.25º	60º N/S	Hourly	2000-present	Sorooshian et al. (2000)
PERSIANN-CCS *	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Cloud Classification System (CCS)	0.04º	60º N/S	Hourly	2003-Present	Hong et al. (2004)
SM2RAIN-ASCAT V1.2 *	Precipitation Estimation from the application of the SM2Rain algorithm to the ASCAT soil moisture data	12.5km	Global	Daily	2007-2019	Brocca et al. (2019)
GPM+SM2RAIN *	Integration of IMERG-E with SM2RAIN-based rainfall estimates derived from three different satellite Soil Moisture products	0.25º	60º N/S	Daily	2007-2018	Massari et al. (2020)
PDIR-Now *	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Dynamic Infrared Rain Rate near real-time (PDIR-Now)	0.04º	60º N/S	Hourly	2000-present	Nguyen et al. (2020)
CHIRPS V2.0 **	Climate Hazards group InfraRed Precipitation with Stations (CHIRPS) V2.0	0.05º	50º N/S	Daily	1981-present	Funk et al. (2015)
CMORPH-CRT V1.0 **	CPC MORPHing technique (CMORPH) bias corrected (CRT) V1.0	0.07º	60º N/S	30 min	1998-2019	Joyce et al. (2004); Xie et al. (2017)
IMERG-F V06 **	Integrated Multi-satellitE Retrievals for GPM (IMERG) final run V06	0.1º	60º N/S	30 min	2000-present	Huffman et al. (2019)

MERRA-2 **	Modern-Era Retrospective Analysis for Research and Applications 2	~0.5º	Global	Hourly	1980-present	Gelaro et al. (2017)
MSWEP V2.2 **	Multi-Source Weighted-Ensemble Precipitation (MSWEP) V2.2	0.1º	Global	3-hourly	1979-present	Beck et al. (2017a, 2019b)
PERSIANN-CDR V1R1 **	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Climate Data Record (CDR) V1R1	0.25º	60º N/S	Daily	1983-present	Ashouri et al. (2015)