Validation of GPM IMERG V05 and V06 Precipitation Products over Iran[✉]

SEYED-MOHAMMAD HOSSEINI-MOGHARI

Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

QIUHONG TANG

Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, and University of Chinese Academy of Sciences, Beijing, China

(Manuscript received 12 November 2019, in final form 19 February 2020)

ABSTRACT

This study attempts to assess the validity of the Integrated Multisatellite Retrievals for Global Precipitation Measurement (IMERG) products across Iran. Six IMERG precipitation products (IPPs) including early, late, and final runs for versions 05 and 06 were compared with precipitation data from 76 synoptic stations on a daily scale for the period from June 2014 to June 2018. According to the results, V05 performed better than V06, particularly in early and late runs. The IPPs overestimate precipitation ranging from 5% to 32%; however, IPPs tended to underestimate (overestimate) the amount of precipitation for wet (dry) areas and precipitation classes higher than 5 mm day^{-1} (less than 5 mm day^{-1}). The probability of detection (POD) in IPPs was almost similar (with a median equal to 0.60), whereas other categorical validation metrics like false alarm ratio (FAR) improved in the final run. Our assessments revealed that the dependency of IPPs to the elevation was low, while the error characteristics of IPPs were strongly dependent on the climate and precipitation intensity. For instance, the systematic error varied between less than 12% in dry regions to more than 60% in wet regions. Also, according to modified Kling–Gupta efficiency (KGE) and relative bias (RBias), the performance of IPPs in winter with the highest KGE (ranging from 0.47 to 0.63) and lowest RBias (ranging from 0% to 16%) was better than other seasons. Further improvement is recommended in the satellite sensors and the precipitation retrieval algorithms to achieve a reliable precipitation source.

1. Introduction

Precipitation is one of the most critical climatic elements that plays a vital role in the water cycle. The understanding of spatiotemporal precipitation variability is essential for many applications such as climate studies, water resources planning and management, as well as environmental monitoring (Dezfooli et al. 2018; Gehne et al. 2016; Kidd and Huffman 2011; Rivera et al. 2018). Generally, precipitation data can be obtained from three sources: ground-based observations, model simulations, and remote sensing observations (Hosseini-Moghari et al. 2018). Ground-based observations

provide the most accurate amount of precipitation locally but generalize their point's spatial representativeness to large spatial extents, especially over poorly gauged areas leading to some random and systematic errors arising from spatial interpolation (Ma et al. 2015; Yong et al. 2016). Although model simulations, for example, European Centre for Medium-Range Weather Forecasts (ECMWF; Persson 2001) and Modern-Era Retrospective Analysis for Research and Application (MERRA; Rienecker et al. 2011), have solved the spatial issue of precipitation data, their accuracy is not reliable (Anagnostopoulos et al. 2010). On the other hand, satellite observations provide a unique opportunity to estimate (near) real-time precipitation globally with promising accuracy (Wolters et al. 2011) that would be beneficial, especially for areas like Iran, where ground-based observations are scarce (Javanmard et al. 2010). Therefore, estimating precipitation from remote sensing observations has become

[✉] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-19-0269.s1>.

Corresponding author: QiuHong Tang, tangqh@igsnr.ac.cn

one of the main approaches to measuring precipitation in the last decades (Tan et al. 2015).

Several satellite-based precipitation products (SPPs) have been developed so far, which are different in terms of the purpose of development, input data sources (e.g., microwave and infrared data), precipitation estimation algorithms, spatiotemporal resolution, spatial coverage, temporal span, and latency. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Sorooshian et al. 2000), the Climate Prediction Center morphing technique (CMORPH; Joyce et al. 2004), Tropical Rainfall Measuring Mission (TRMM; Huffman et al. 2007), the Global Precipitation Measurement (GPM; Hou et al. 2014), and PERSIANN–Climate Data Record (PERSIANN-CDR; Ashouri et al. 2015) are examples of these products. Due to the availability of SPPs and limitations in ground-based precipitation data, attention was drawn to consider SPPs as an alternative to ground-based observations (Alijanian et al. 2017). Moreover, applications of SPPs in a broad range of studies indicated their noteworthy potentials in various fields, for example, extreme precipitation analysis, drought monitoring, hydrological modeling, etc. (Camici et al. 2018; Hazra et al. 2019; Lai et al. 2019; Maggioni and Massari 2018; Wang et al. 2019; Yong et al. 2016; Zhang et al. 2019; Zhu et al. 2019; Zorzetto and Marani 2019; Zubieta et al. 2019).

Despite improvements in satellite-based precipitation estimation, there is no guarantee for the validity of their data in all areas as their performance may vary from region to region or from season to season. For instance, Darand et al. (2017) indicated that the TRMM Multisatellite Precipitation Analysis (TMPA) showed a better performance in western parts of Iran and wet periods compared to the other areas and dry periods. As another example, a high correlation was reported between CMORPH products and observed data over China (Guo et al. 2016; Yang et al. 2016). In contrast, the mentioned correlation was low in Iran (Alijanian et al. 2017; Katirai-Boroujerdy et al. 2013) and also in Italy (Duan et al. 2016). Therefore, assessing the validity of SPPs and understanding of SPPs' error characteristics are vital to comprehend their performance in a given area before using them in operational applications. These preliminary assessments are beneficial to progress in the SPPs' precipitation estimation algorithm and satellite instruments (Tan and Santo 2018).

The GPM satellite has benefited from these assessments in both satellite instruments and precipitation estimation algorithms. The frequency range of GPM microwave imager (13 channels) is 4 channels higher than TRMM (9 channels), and the spaceborne precipitation radar

instruments on board GPM are more advanced than TRMM (Tan and Santo 2018). Also, spatiotemporal resolution and spatial coverage in GPM have been improved in comparison with TRMM. Currently, the Integrated Multisatellite Retrievals for GPM (IMERG) version 6 as the newest GPM precipitation product, with a half-hourly temporal resolution and $0.1^\circ \times 0.1^\circ$ spatial resolution has a full coverage of the globe, whereas TRMM products cover a smaller latitude band (50°N – 50°S) with a lower spatial ($0.25^\circ \times 0.25^\circ$) and temporal (3-hourly) resolution (Huffman et al. 2007). Furthermore, in terms of algorithm progress, IMERG utilizes a combination of three SPPs features, including PERSIANN, CMORPH, and TMPA (Anjum et al. 2018; Prakash 2019). It is worth mentioning that PERSIANN estimates the precipitation based on infrared brightness temperature image (as input) and artificial neural network (as a model) (Ashouri et al. 2015), while CMORPH is mainly based on microwave data and only uses infrared data for transporting the microwave-based precipitation features within periods that microwave data are not available at a location (Joyce et al. 2004). Because of these improvements, IMERG precipitation products (IPPs) appeared to be better than TMPA and other SPPs in many regions, for example, India (Prakash et al. 2018), Pakistan (Anjum et al. 2018), Brazil (Rozante et al. 2018), and East Asia (Lee et al. 2019).

Although the IPP has been evaluated for many areas, at present, only a few studies have attempted to assess the reliability of IMERG products for Iran; one study compared IMERG, TMPA-3B42, and ERA-Interim precipitation products for four provinces in Iran from March 2014 to February 2015 (Sharifi et al. 2016). The results showed that all products underestimated precipitation in the studied areas, but there was a slight underestimation for IMERG. In another study, Khodadoust Siuki et al. (2017) evaluated IMERG and TMPA products in a province in the northeast of Iran (Khorasan Razavi Province) from March to December 2014. Their study illustrated that both products underestimated rainfall, while IMERG had a better correlation with in situ data. Recently, Aslami et al. (2019) analyzed the daily precipitation from 17 gauges between 1 January 2016 and 21 October 2017 in Ardabil Province (northwestern Iran). The analysis showed that although IPP values were relatively close to observations, their accuracy was not high, for example, in more than 96% of selected pixels, IPP tended to overestimate the precipitation by 24% in annual scale. To the best of our knowledge, there is no assessment of IPPs over all of Iran, and its latest version (version 6) has not been considered at all. Considering the elevation variations, different climates, and non-uniform spatiotemporal distribution of precipitation

FIG. 1. Map of the study area with the location of synoptic stations and their climates. Some stations in the south are located on islands.

over Iran, a primary assessment of IPP over Iran can be a useful procedure to reveal the error characteristics of this SPP. Therefore, in this study, we investigated the following threefold objectives:

- (i) How valid are IPPs over Iran?
- (ii) Are there any improvements toward the accuracy of IPPs version 6 (V06) runs compared to the version 5 (V05) runs?
- (iii) How do IPPs perform in different elevations, precipitation intensities, climatic zones, and seasons across Iran?

To this end, precipitation data from 76 synoptic stations and three runs of the IMERG V05 and V06 products were utilized between 1 June 2014 and 30 June 2018. In sections 2 and 3, the study area and employed data were described in more detail, respectively. Section 4 represented the methodology of the study. The results of the assessments were shown in section 5. Section 6 is the discussion section. Finally, conclusions are presented in section 7.

2. Study area

Iran was considered as the case study in this paper (Fig. 1). Iran, with an area of 1 648 000 km², is located in southwest Asia between longitudes 44° and 64°E and between latitudes 25° and 40°N (Rahimi et al. 2013). Although the majority of the country has an arid and semiarid climate, Iran's climate varies from extra arid to very humid (Alijani et al. 2008). Iran contains two high

mountain chains, namely the Alborz mountain chain in the north and the Zagros mountain chain in the west, and two vast deserts (the Kavir and Lut Deserts). As a result, elevations vary between −28 and 5610 m MSL across Iran (Fig. 1). The average annual precipitation of Iran is 241 mm (Tabari and Talaei 2011) and ranges from less than 50 mm in deserts to about 1800 mm in the north parts (Yazdanpanah et al. 2017).

3. Datasets

a. Observed precipitation dataset

High-quality daily precipitation data from 76 key synoptic stations across Iran (Fig. 1) from 1 June 2014 to 30 June 2018 were used in this study. For a double check, the quality of data has been checked in terms of missing data and the amount of precipitation. The number of days without data was insignificant (maximum 1.8% in Siri Island station; for details please refer to Table S1 in the online supplemental material); hence, no reconstruction has been done to deal with missing data (missing data were removed from the calculations). To verify the amount of precipitation, we compared the summation of daily data in each month with monthly quality controlled precipitation data that were released by the Islamic Republic of Iran Meteorological Organization (IRIMO). There was no difference between the two datasets. The observed data were downloaded through <http://irimo.ir/eng/wd/720-Products-Services.html>. Zabol and Anzali stations with annual precipitation 35 and 1757 mm

TABLE 1. Characteristics of IMERG precipitation products.

Products	Spatial			Temporal
IMERG Early V05 (V05E)	$0.1^{\circ} \times 0.1^{\circ}$	Global	30 min	12 Mar 2014–1 May 2019
IMERG Late V05 (V05L)	$0.1^{\circ} \times 0.1^{\circ}$	Global	30 min	12 Mar 2014–26 Aug 2019
IMERG Final V05 (V05F)	$0.1^{\circ} \times 0.1^{\circ}$	Global	30 min	12 Mar 2014–30 Jun 2018
IMERG Early V06 (V06E)	$0.1^{\circ} \times 0.1^{\circ}$	Global	30 min	1 Jun 2000–present
IMERG Late V06 (V06L)	$0.1^{\circ} \times 0.1^{\circ}$	Global	30 min	1 Jun 2000–present
IMERG Final V06 (V06F)	$0.1^{\circ} \times 0.1^{\circ}$	Global	30 min	1 Jun 2000–present

(between 1 June 2014 and 31 May 2018) have the lowest and highest amount of precipitation among the studied stations, respectively. Also, Anzali and Baft stations are the lowest and the highest stations with an elevation of -23.6 and 2280 m MSL, respectively. The main characteristics of the stations are presented in the online supplemental material (Table S1).

b. GPM IMERG precipitation products

GPM is a joint mission between the National Aeronautics Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) that was launched on 27 February 2014, to provide high spatiotemporal precipitation data globally. GPM is responsible for extending the TRMM mission to create the next generation of Earth's precipitation. In addition to one core observatory, GPM also consists of approximately 10 constellation satellites (Fang et al. 2019). The GPM Microwave Imager (GMI) and the Dual-Frequency Precipitation Radar (DPR) are the two primary sensors of GPM satellite that are used to estimate precipitation type and intensity, as well as to detect the internal structure of storms under and within clouds, respectively (Tan and Santo 2018).

The U.S. GPM team using the IMERG algorithm, as a unified algorithm, generates the GPM level-3 precipitation products, that is, IPP. Three types of IMERG products are available in each version, namely, early (E), late (L), and final (F) runs with a latency of 4 h, 12–14 h, and 3.5 months, respectively (Huffman et al. 2015; Tan et al. 2019). IMERG early and late runs can be used for real-time applications such as flood monitoring and irrigation management, while the final run is a research-grade product. All runs of IPPs are currently available with 30-min temporal resolution and $0.1^{\circ} \times 0.1^{\circ}$ spatial resolution over the fully global domain. Some changes in the new version (V06) of IMERG products are as follows: 1) the “displacement vectors,” which are used for time interpolation, are calculated based on Modern-Era Retrospective Reanalysis 2 (MERRA-2) and Goddard Earth Observing System model (GEOS) Forward Processing (FP) data instead of the infrared data that were used in V05; 2) the estimation of Sounder for Atmospheric Profiling of Humidity in the Intertropics

by Radiometry (SAPHIR) is used for the first time in V06; and 3) in IMERG V06, TRMM-based calibrations are applied for the first 2.5 months of the GPM era to let the GPM-based calibrations spin up before those are employed, while in V05 GPM-based calibrations were started immediately (Huffman et al. 2015).

In this study, we used early, late, and final runs of V05 and V06 for the period of 1 June 2014–30 June 2018 (the common period in the two versions and all runs). Table 1 shows the main characteristics of IPPs used in this study. It should be noted that reporting time of daily precipitation data in synoptic stations is 0600 UTC (WMO 2009). Therefore, if we aggregate data from 0000 to 2359 UTC, a quarter of daily precipitation accumulations would be assigned to the wrong day. Hence, to produce the daily IPPs to coincide with daily data in synoptic stations, we aggregated half-hourly data from 0600:00 UTC on a given day to 0559:59 UTC on the next day. The IMERG products are available at <https://disc.gsfc.nasa.gov/>.

4. Methodology

Considering the high resolution of IPPs and significant uncertainties arising from interpolation of the daily data of 76 synoptic stations across Iran (due to the low-density of stations), we applied a point-to-pixel approach for evaluations in this study like many previous studies, for example, Xu et al. (2017), Rivera et al. (2018), Wei et al. (2018a), and Satgé et al. (2018). To assess the validity of IMERG products over Iran, we considered continuous validation metrics such as modified Kling–Gupta efficiency (KGE), relative bias (RBias), correlation coefficient (CC), and root-mean-square error (RMSE); as well as categorical validation metrics such as probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and Heidke skill score (HSS). Moreover, an error decomposition into systematic error (SE) and random error (RE) components was applied based on the Willmott approach (Willmott 1981). The equations of these metrics were presented in Table 2.

KGE, introduced in Gupta et al. (2009) and modified by Kling et al. (2012), is a performance metric that considers the distance between mean and variance of

TABLE 2. List of the assessment metrics to quantify the validity of the IPPs. Note that $\overline{P^s}$ and $\overline{P^g}$ are the mean of satellite and gauge-based precipitation, respectively, and σ^s and σ^g are the standard deviation of satellite and gauge-based precipitation, respectively. Also, n is the total number of gauge observation data (or IMERG data), P_i^s and P_i^g are i th satellite and gauge-based precipitation, respectively. The terms a and b are slope and intercept of a linear regression fitted on gauge precipitation data, respectively. The term H represents a hit (number of precipitation events that was correctly detected by the satellite), M represents a miss (number of precipitation events that was not detected by the satellite), F represents a false alarm (number of precipitation events that was detected by the satellite but did not occur), and C represents a correct negative (number of precipitation events that was not detected by the satellite and did not occur).

No.	Metric	Equation	Perfect value	Unit
1	Kling–Gupta Efficiency (KGE)	$\text{KGE} = 1 - \sqrt{(\text{CC} - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$ $\beta = \frac{\overline{P^s}}{\overline{P^g}}, \quad \gamma = \frac{(\sigma^s/\overline{P^s})}{(\sigma^g/\overline{P^g})}$	1	—
2	Relative bias (RBias)	$\text{RBias} = \frac{\sum_{i=1}^n (P_i^s - P_i^g)}{\sum_{i=1}^n P_i^g} \times 100$	0	%
3	Correlation coefficient (CC)	$\text{CC} = \frac{\sum_{i=1}^n (P_i^g - \overline{P^g})(P_i^s - \overline{P^s})}{\sqrt{\sum_{i=1}^n (P_i^g - \overline{P^g})^2} \sqrt{\sum_{i=1}^n (P_i^s - \overline{P^s})^2}}$	1	—
4	Root-mean-square error (RMSE)	$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (P_i^s - P_i^g)^2}{n}}$	0	mm
5	Systematic error (SE)	$\text{SE} = \frac{\sum_{i=1}^n (P_i^g - P_i^s)^2}{n} \Bigg \text{MSE} \times 100$	0	%
6	Random error (RE)	$\text{RE} = \frac{\sum_{i=1}^n (P_i^s - P_i^g)^2}{n} \Bigg \text{MSE} \times 100$ $\text{MSE} = \frac{\sum_{i=1}^n (P_i^s - P_i^g)^2}{n}, \quad P_i^g = a \times P_i^s + b$	0	%
7	Probability of detection (POD)	$\text{POD} = \frac{H}{H + M}$	1	—
8	False alarm ratio (FAR)	$\text{FAR} = \frac{F}{H + F}$	0	—
9	Critical success index (CSI)	$\text{CSI} = \frac{H}{H + M + F}$	1	—
10	Heidke skill score (HSS)	$\text{HSS} = \frac{2(H \times C - F \times M)}{(H + F) \times (F + C) + (H + M) \times (M + C)}$	1	—

observed and estimated time series as well as their correlation as a criterion for evaluation (see Table 2). RBias indicates an overestimation or underestimation of total precipitation by SPPs. CC, ranging from -1 to 1 , is computed to determine the linear correlation between the SPPs and observed precipitation. RMSE is used to measure the mean error magnitude between the SPPs and the gauge-based precipitation (Rivera et al. 2018).

POD ranges from 0 to 1 (the perfect score) and is utilized to estimate the hit rate. POD denotes the ratio of the number of precipitation events that are correctly detected by the SPPs to the total number of actual events (Sharifi et al. 2018). FAR represents the fraction of falsely detected events (Xu et al. 2017). With values ranging from 0 (the perfect value) to 1 , a FAR of 0

indicates there is no wrong detected event, and a value of 1 shows that all detected events by the SPPs are a false alarm (Ghajarnia et al. 2015). The CSI as a function of POD and FAR, with values between 0 and 1 (the perfect value), measures the overall performance of the SPPs to recognize the correct detection of precipitation events (Wei et al. 2018a). HSS, which ranges from minus infinity to 1 (the perfect value), is computed to estimate the accuracy after eliminating those detections that are just detected correctly because of the random chance (Zappa 2008). A negative value of HSS shows that a random estimation is better than the SPP, and HSS equal to 0 means that SPP has no skill, while an HSS value of 1 demonstrates the perfect performance of SPP (Diem et al. 2014). For calculating the categorical

TABLE 3. Elevation categories of the studied stations over Iran (Khalili and Rahimi 2014).

Elevation classes (m MSL)	No. of stations
<500	26
500–1000	7
1000–1500	30
1500–2000	9
>2000	4

validation metrics, since the tiny precipitation events are not important from a meteorological and hydrological point of view, it is common to consider a threshold to eliminate their effect on the results of assessments. Therefore, the rain/no-rain threshold was assumed to be 1 mm day^{-1} , as reported by Alijanian et al. (2017) and Satgé et al. (2018).

We also decomposed mean squared error (MSE) of IPPs to their systematic and random error components, following the method of AghaKouchak et al. (2012). The random error mainly depends on finite sampling and estimation algorithm, while systematic problems mainly result in biases such as the inclusion of rain gauges that is only available on land (Huffman 1997). Therefore, understanding the role of random and systematic components is essential to develop the next-generation bias removal techniques, and to improve precipitation retrieval algorithms as well as some other applications (AghaKouchak et al. 2012; Maggioni et al. 2016). Decomposing the SPPs error to the systematic and random parts can be done by using Eqs. (5) and (6) from Habib et al. (2009).

In addition to a general assessment, we investigated the impacts of elevation, precipitation intensity, and climatic zone on error characteristics of IPPs over Iran. To assess the effects of elevation on the performance of IPPs, we divided all studied stations into five elevation categories, according to Khalili and Rahimi (2014), that are presented in Table 3. Assessing the effects of elevation on IPPs' performance was done to illustrate the sensitivity of the IMERG retrieval algorithm to the elevation. Moreover, we classified daily precipitation intensities into four categories based on Alijanian et al. (2017) who classified the precipitation intensity over Iran. The precipitation intensity categories were shown in Table 4. Besides, to consider the effect of climatic zones, we classified stations to different climatic zones (Fig. 1) based on the De Martonne classification (De Martonne 1926) that was reported for Iran by Khalili and Rahimi (2018) and Rahimi et al. (2013). In short, in the De Martonne classification, the climatic zones are identified based on an aridity index (AI) that would be calculated using $AI = P/(T + 10)$, where P is the long-

TABLE 4. Precipitation intensity categories over Iran (Alijanian et al. 2017).

Precipitation classes	Amount of precipitation (mm day^{-1})
No/tiny	Rain < 1
Light and low moderate	$1 \leq \text{rain} < 5$
High moderate and low heavy	$5 \leq \text{rain} < 20$
High heavy and violent	Rain ≥ 20

term average of annual precipitation (mm) and T is the long-term average of mean annual temperature ($^{\circ}\text{C}$). The climatic zones were presented in Table 5. As the last step, the assessment of IPPs was done in different seasons. To this end, daily precipitation data in the months December–February, March–May, June–August, and September–November correspond to winter, spring, summer, and autumn, respectively.

5. Results

a. General analysis

Figure 2 indicates the scatterplots of six IPPs for the studied stations from 1 June 2014 to 30 June 2018 (number of points is 112 982). From the figure, there are no significant differences between the early and late runs of IPPs, whereas high overestimated precipitation values are reduced in the final run. In all products, precipitation amounts greater than 100 mm are significantly underestimated, but many precipitation events with less than 50 mm are overestimated, particularly in early and late runs. The slope of the blue lines (regression lines) is always less than one (on average about 0.5), indicating that gauge data are greater than IPPs. It should be noted that in the arid and semiarid areas like Iran, there is no precipitation on the majority of days (see Fig. 3); as a result, the linear equation mainly refers to no rainy days. Moreover, a small positive intercept illustrates that the IPPs have estimated an insignificant amount of precipitation on dry days. As a result, results in the probability of daily precipitation less than 1 mm day^{-1} for IPPs approximately equals to the probability of no rainy days (less than 0.1 mm day^{-1}) in the observed dataset (see Fig. 3). Therefore, as a simple adjustment, the rain/no-rain threshold was assumed to be equal to 0.1 and 1 mm day^{-1} for the observed data and the IPPs, respectively. Figure 3 shows that the probability of precipitation less than 0.1 mm in gauges is more than 0.85, while this lies between 0.75 and 0.79 for IPPs. Although the empirical cumulative distribution functions (eCDFs) for all IPPs indicate a similar behavior, the precipitation amounts are not matched to the observed eCDF.

TABLE 5. Climate zones of the studied stations based on De Martonne classification (Rahimi et al. 2013; Khalili and Rahimi 2018).

Climatic zone	No. of stations	Aridity index
Extra arid	19	$AI < 5$
Arid	14	$5 \leq AI < 10$
Semiarid	33	$10 \leq AI < 20$
Mediterranean	2	$20 \leq AI < 24$
Subhumid	2	$24 \leq AI < 28$
Humid	2	$28 \leq AI < 35$
Per-humid A	3	$35 \leq AI < 55$
Per-humid B	1	$AI \geq 55$

The eCDF of IMERG V06F can be seen in front of IMERG V05F for those precipitation events less than 1 mm, while both eCDFs almost converged together for precipitation amounts higher than 1 mm. However, eCDF of the observed precipitation lies behind IPPs, which is an indication of overestimation in all IPPs.

Figure 4 shows box-and-whisker plots of annual precipitation along with precipitation regimes for the studied stations based on gauge observations and IPPs. The mean annual precipitation of each station

was plotted as the blue point in Fig. 4a. Based on Fig. 4a, the high annual precipitation (more than 1000 mm yr^{-1}) has been significantly underestimated, but the low precipitation has been overestimated. It should be noted that the total number of annual precipitation amounts more than 1000 mm yr^{-1} is equal to 16 during the study period. The final runs have experienced improvements in the overestimations and underestimations but have not yet performed well. However, the mean of annual precipitation over all stations was estimated with acceptable accuracy in final runs. The mean annual precipitation over all stations was estimated by IMERG V05E, V05L, V05F, V06E, V06L, and V06F equal to 338, 324, 305, 383, 382, and 312 mm, respectively, while the observed one was 291 mm (mean from 1 June 2014 to 31 May 2018). With respect to bias, these results indicate the better performance for V05 than V06, specifically in early and late runs. Based on Fig. 4b, the worst performance of the IPPs can be seen between February and May, where IMERG V06E and V06L, with a similar dynamic with V05, overestimate precipitation up to near 100% in March. The regimes of precipitation in

FIG. 2. Scatterplot of daily precipitation for observed precipitation vs (a) IMERG V05E, (b) IMERG V05L, (c) IMERG V05F, (d) IMERG V06E, (e) IMERG V06L, and (f) IMERG V06F.

FIG. 3. Empirical cumulative distribution function of daily precipitation for observed precipitation and the IPPs. The x axis has a logarithmic scale.

final runs are closer to the observed one; however, those still overestimate the precipitation from January to June and underestimate that between August and November.

Box-and-whisker plots of assessment metrics for the studied stations on a daily scale were shown in Fig. 5. All assessment metrics (except the POD) have been significantly improved in the final run. For instance, the median of KGE values for V05E, V05L, V05F, V06E, V06L, and V06F become equal to 0.27, 0.30, 0.50, 0.15, 0.20, and 0.51, respectively. With respect to KGE, V06E, and V06L perform worse than the same run in V05. The median of CC values near 0.60 for early and late runs

and 0.70 for final runs indicate that IMERG products follow the behavior of observed data with relatively acceptable accuracy. The positive bias in the most stations (more than three quarters of the stations) shows IPPs overestimate the amount of precipitation resulted in the medians of RMSE values vary from 2.2 mm (for V05F) to 3.4 mm (for V06L). Assessing categorical validation metrics show there is no improvement in the median of POD in different IPPs; it equals 0.60 in different versions and runs. On the other hand, FAR, CSI, and HSS have been systematically improved from early runs to final runs. Also, based on three last-mentioned metrics, V05 has a better performance in comparison with V06 in the same run. For instance, the median of FAR values for early, late, and final runs in V05 equal to 0.38, 0.32, and 0.30, respectively, while 0.39, 0.36, and 0.32 are obtained for V06.

Figures 6 and 7 illustrate the spatial distribution of continuous validation metrics and categorical validation metrics, respectively. From Fig. 6, in early and late runs, KGE values are almost negative in the south of Iran, which are probably affected by large RBias in this area. Generally, KGE values are significantly improved in the final runs. The lowest CCs and highest negative RBias can be seen in the south of the Caspian Sea, suggesting that the IPPs neither estimate rainfall dynamic well nor its amount resulted in RMSE more than 10 mm in this region. In early and late runs for V06, RBias is worse than V05 somehow in the south and west of Iran, the underestimation and overestimation have been increased, respectively. Overall, V06 tends to overestimate precipitation more than V05. Figure 6 clearly reveals that the RMSE is significantly improved in final runs in all regions except south of the Caspian Sea. According to the RSME, V05 also performs better than V06 in most areas,

FIG. 4. Box-and-whisker plots of (a) annual precipitation and (b) monthly precipitation regimes for observed precipitation and the IPPs for the studied stations from 1 Jun 2014 to 31 May 2018.

FIG. 5. Box-and-whisker plots of assessment metrics for the IPPs for the studied stations on a daily scale.

particularly in early and late runs. However, there is a slight difference in the final runs.

Figure 7 shows POD values in the west and south of Iran are higher than 0.7, while in the south of the Caspian Sea and for early and late runs are less than 0.50, these findings indicate that half of the precipitation events cannot be captured by the satellite in this region. Contrary to the previous metric, FAR values have been improved in numerous stations in late runs in comparison with early runs. However, the median of FAR is 0.30 for the best situation (V05F) showing that more than 30% of detected precipitation events are a false alarm in the fifty percent of the stations. Median of CSI values in all runs and versions are less than 0.46 which means that the satellite correctly detected less than half of the precipitation events (observed and/or predicted). HSS metric, as same as other metrics, improved over the most parts of Iran in the final run. However, this improvement in the south of the Caspian Sea is marginal. Medians of HSS values ranging from 0.51 to 0.58 show the capability of IPPs to estimate precipitation in the majority of the stations while their skills are not close to the perfect one ($HSS = 1$). In all stations, HSS values are greater than 0.26, suggesting that IPPs can capture the precipitation events.

b. Evaluation of elevation impact

The performance of IPPs was investigated in different elevation classes (see Table 2). Assessment metrics for

different elevation classes were presented in Table 6. From the table, RBias values for the elevations between 500 and 1000 m are enormous, especially in early and late runs. It leads to the lowest values of KGE, while the CC values are better than the elevations below than 500 m. Nevertheless, the performance of IPPs has been significantly improved in final runs. The maximum RMSE values can be grasped among the elevations below 500 m that are near twice the rest. It should be noted that almost all stations with high precipitation (wet regions) are located below 500 m (see Fig. 1). Evaluating the decomposed errors indicates that the contribution of SE is not significant (maximum 21%) in all elevation classes except for the below 500 m. In V05 for the elevations below 500 m, SEs can be considered as almost half of the errors. However, in V06, the contribution of SEs is reduced in all elevation classes (except for early and late runs in the elevations 1500–2000 m). Based on Table 6, the lowest POD values are related to the elevations below 500 m, while the lowest FAR and the highest values CSI and HSS can be seen above 2000 m. However, a significant relationship cannot be found between the elevation and categorical validation metrics.

The scatterplots of each evaluation metric against elevation were drawn in Fig. 8. As can be seen in Fig. 8, each point indicates one station, and each color corresponds to one IPP. Generally, Fig. 8 shows that the

FIG. 6. Spatial distribution of KGE, CC, RBias, and RMSE for the IPPs for the studied stations on a daily scale.

dependence on elevation seems quite low. However, the lowest correlation values and the biggest RMSE, SE, and negative RBias are observed in negative elevations where the highest amount of precipitation falls (south of

the Caspian Sea). Nevertheless, since there is no significant relationship between metrics and elevation, another factor should be responsible for the poor performance of IPPs in negative elevations, for example,

FIG. 7. Spatial distribution of POD, FAR, CSI, and HSS for the IPPs for the studied stations on a daily scale.

climatic conditions. For further assessment, the characteristics of the linear relationship between assessment metrics and elevations were considered (see Table S2). Based on this table, no significant relationship is seen in

most cases. With respect to categorical validation metrics, there is no statistically significant relationship between categorical validation metrics and elevations. However, a significant inverse relationship between

TABLE 6. Calculated assessment metrics for the IPPs in different elevation classes on a daily scale.

Elevation (m MSL)	Metrics	V05E	V05L	V05F	V06E	V06L	V06F
<500	KGE	0.34	0.36	0.49	0.35	0.36	0.49
	CC	0.36	0.37	0.53	0.36	0.37	0.51
	RBias (%)	−10	−12	−15	11	11	−12
	RMSE (mm)	6.47	6.42	5.52	7.25	7.35	5.74
	SE (%)	47	46	46	28	26	39
	RE (%)	53	54	54	72	74	61
	POD	0.52	0.52	0.54	0.52	0.51	0.54
	FAR	0.41	0.35	0.30	0.43	0.39	0.32
	CSI	0.38	0.40	0.44	0.37	0.39	0.43
	HSS	0.49	0.52	0.56	0.47	0.49	0.54
500–1000	KGE	0.02	0.09	0.47	−0.21	−0.15	0.47
	CC	0.58	0.62	0.74	0.56	0.55	0.73
	RBias (%)	84	78	37	111	105	39
	RMSE (mm)	2.43	2.33	1.64	3.19	3.28	1.71
	SE (%)	6	5	12	2	1	8
	RE (%)	94	95	88	98	99	92
	POD	0.61	0.61	0.60	0.60	0.61	0.60
	FAR	0.48	0.43	0.36	0.49	0.44	0.37
	CSI	0.39	0.42	0.44	0.38	0.41	0.44
	HSS	0.51	0.55	0.58	0.50	0.54	0.57
1000–1500	KGE	0.47	0.53	0.62	0.41	0.42	0.61
	CC	0.64	0.66	0.74	0.65	0.66	0.73
	RBias (%)	38	32	25	46	47	24
	RMSE (mm)	3.25	3.09	2.42	3.34	3.47	2.48
	SE (%)	3	3	8	2	1	8
	RE (%)	97	97	92	98	99	92
	POD	0.57	0.57	0.59	0.56	0.55	0.57
	FAR	0.39	0.35	0.31	0.40	0.39	0.34
	CSI	0.42	0.44	0.47	0.41	0.41	0.44
	HSS	0.52	0.55	0.58	0.51	0.51	0.55
1500–2000	KGE	0.47	0.51	0.63	0.34	0.32	0.59
	CC	0.66	0.66	0.76	0.64	0.63	0.75
	RBias (%)	40	34	22	51	54	27
	RMSE (mm)	3.19	3.03	2.26	3.11	3.27	2.30
	SE (%)	3	5	12	7	6	11
	RE (%)	97	95	88	93	94	89
	POD	0.59	0.58	0.60	0.61	0.60	0.60
	FAR	0.38	0.36	0.31	0.43	0.44	0.36
	CSI	0.43	0.44	0.47	0.42	0.41	0.45
	HSS	0.53	0.54	0.58	0.51	0.50	0.55
>2000	KGE	0.53	0.64	0.71	0.44	0.48	0.72
	CC	0.67	0.69	0.78	0.63	0.64	0.78
	RBias (%)	31	14	9	41	37	10
	RMSE (mm)	2.99	2.70	2.15	3.51	3.54	2.19
	SE (%)	7	15	21	4	2	18
	RE (%)	93	85	79	96	98	82
	POD	0.61	0.60	0.61	0.61	0.59	0.60
	FAR	0.36	0.32	0.28	0.37	0.34	0.29
	CSI	0.45	0.47	0.49	0.45	0.45	0.48
	HSS	0.57	0.59	0.62	0.57	0.57	0.60

RMSE and elevation (CC ranging from −0.39 to −0.57) was found which is mostly because of the large error in the south of the Caspian Sea where the elevation is below sea level. Assessing continuous validation metrics in various elevations shows V05 outperforms V06 in most cases. For instance, in Fig. 8, RBias of V06 was usually located above RBias V05 in all elevations, resulted in

the median of RBias for V05E, V05L, V05F, V06E, V06L, and V06F were equal to 52, 45, 25, 62, 60, and 30, respectively.

c. Evaluation of precipitation intensity impact

Figure 9 demonstrates the frequency of daily precipitation for four precipitation classes (see Table 3).

FIG. 8. Relationships between calculated assessment metrics for IMERG products and elevations on a daily scale.

According to the figure, the frequency of days with the precipitation below 1 mm day^{-1} in all IPPs is less than the observed one, whereas the opposite is true for classes 1–5 and 5–20 mm day^{-1} . For instance, the frequency of precipitation less than 1 mm day^{-1} in the observed dataset is

equal to 91% and for V05F and V06F are 89% and 88%, respectively; while in the class 1–5 mm day^{-1} , the frequency of observed data is 5% and for both V05F and V06F is 7%. The frequency of precipitation with more than 20 mm day^{-1} in observed, V05E, and V05L datasets

FIG. 9. Frequency of daily precipitation for observed precipitation and the IPPs in different precipitation classes for the studied stations.

is equal to 0.9%, while for V06E and V06L is 1% and 1.1%, respectively. Also, in both final runs the frequency of daily precipitation with more than 20 mm day⁻¹ is 0.8%.

Assessment metrics for different precipitation classes were presented in Table 7. In this table, FAR, CSI, and HSS were not provided because if all days in the observed dataset are rainy, the value of FAR and HSS will be equal to zero and CSI equal to POD (except the “All” class that includes no rainy days). Also, POD was not calculated for the class less than 1 mm day⁻¹ due to the rain/no-rain threshold. From Table 7, generally, the continuous validation metrics in each class are worse than the situation that all data were considered together, particularly in terms of CC, RBias, and KGE. It means that some errors such as overestimation and underestimation can be decreased when the whole of the time series are considered. Therefore, applying IPPs data at the event scale is more error prone than using the entire time series. As expected, POD values are improved when precipitation rates are increased. Also, SE for the precipitation less than 20 mm day⁻¹ is negligible (maximum 6%), which is significantly increased for the precipitation more than 20 mm day⁻¹ (maximum 61%). In general, for precipitation class exceeding 20 mm day⁻¹, a large POD value suggests that IPPs are capable of capturing the occurrence of precipitation events, while low CC values and their poor performance based on RBias and RMSE indicate that IPPs could neither follow the dynamic of precipitation nor estimate well the amount of precipitation. Overall, IPPs performance is variable in different precipitation intensities.

d. Evaluation of climatic zone impact

For evaluating the effects of climate (see Fig. 1 and Table 4), assessment metrics were calculated for each

climatic zone, which are presented in Table 8. In this table, results for stations were located in extra arid and arid regions, as well as per-humid A and B have been merged due to a similar pattern. However, the results for each climatic zone were presented separately in the online supplemental material (Table S3). From the table, in subhumid and humid zones, there is no significant difference between final runs and early or late runs, especially in V05. In the extra arid, arid, semiarid, and Mediterranean climates, all IPPs overestimate the precipitation, while in humid and per-humid A and B climates always suffer from underestimation. Generally, the CC values do not immensely vary from extra arid to humid climates, but the highest CC values are found in the Mediterranean climate that reach to 0.81 (0.80) in V05F (V06F). Also, the lowest values of CC can be found in per-humid A and B climates ranging from 0.29 to 0.45. SEs from extra arid to Mediterranean climates are negligible, whereas from subhumid to per-humid A and B SEs are constantly increased (to 87%). With respect to POD, CSI, and HSS the worst results are seen in per-humid A and B climates. However, the FAR scores in extra arid and arid climates are worse than other climatic zones. The highest POD, CSI, and HSS values are related to the Mediterranean climate ranging [0.68, 0.70], [0.49, 0.56], and [0.58, 66], respectively. In sum, according to Table 8, it can be argued that IPPs performance relates to climate conditions.

Due to the aforementioned results, the performance of IPPs depends on precipitation intensities and climatic zones. Therefore, we have considered the joint effects of precipitation event intensities and climatic zones in Fig. 10. As can be seen in Fig. 10, the number of days per station (NDs) with precipitation less than 1 mm day⁻¹

TABLE 7. Calculated assessment metrics for the IPPs in different precipitation classes on a daily scale.

Precipitation (mm day ⁻¹)	Metrics	EV05	LV05	FV05	EV06	LV06	FV06
All	KGE	0.42	0.45	0.57	0.36	0.37	0.56
	CC	0.47	0.48	0.61	0.46	0.47	0.59
	RBias (%)	17	12	5	32	32	7
	RMSE (mm)	4.55	4.45	3.72	5.01	5.11	3.85
	SE (%)	28	28	33	17	14	28
	RE (%)	72	72	67	83	86	72
	POD	0.56	0.56	0.58	0.56	0.55	0.57
<1	KGE	-15.34	-13.34	-10.28	-18.13	-16.95	-11.28
	CC	0.22	0.23	0.25	0.21	0.21	0.24
	RBias (%)	1632	1432	1126	1911	1793	1225
	RMSE (mm)	1.62	1.50	1.15	1.91	1.91	1.20
	SE (%)	6	6	6	6	6	6
	RE (%)	94	94	94	94	94	94
	POD	0.55	0.55	0.58	0.55	0.54	0.56
1-5	KGE	-2.04	-1.90	-1.28	-2.47	-2.45	-1.51
	CC	0.12	0.13	0.16	0.10	0.11	0.14
	RBias (%)	50	46	31	67	67	30
	RMSE (mm)	6.64	6.20	4.63	8.22	8.19	4.92
	SE (%)	4	4	4	4	4	3
	RE (%)	96	96	96	96	96	97
	POD	0.55	0.55	0.58	0.55	0.54	0.56
5-20	KGE	-1.76	-1.71	-0.95	-1.76	-1.82	-1.15
	CC	0.22	0.23	0.30	0.20	0.21	0.30
	RBias (%)	-13	-13	-14	-7	-4	-16
	RMSE (mm)	12.28	12.13	9.25	13.22	13.79	9.76
	SE (%)	2	2	4	1	1	4
	RE (%)	98	98	96	99	99	96
	POD	0.75	0.76	0.79	0.75	0.75	0.79
≥20	KGE	-0.48	-0.46	-0.12	-0.58	-0.57	-0.19
	CC	0.08	0.08	0.21	0.10	0.09	0.21
	RBias (%)	-51	-50	-45	-42	-38	-42
	RMSE (mm)	35.92	35.70	31.78	38.27	39.10	32.64
	SE (%)	61	61	60	43	38	53
	RE (%)	39	39	40	57	62	47
	POD	0.79	0.80	0.83	0.79	0.79	0.83

was underestimated by all IPPs from extra arid to subhumid climatic zones where the opposite is true for pre-humid A and B climates. NDs with precipitation between 1 and 5 mm day⁻¹ are overestimated in all climates. However, for precipitation events that are higher than 5 mm day⁻¹, NDs in humid and per-humid A and B are significantly underestimated and vice versa for the rest of climatic zones (except for the precipitation events higher than 20 mm day⁻¹ in the Mediterranean climate). These results suggest that precipitation volume is overestimated and underestimated in dry and wet regions, respectively (see Fig. 4a). To sum up, high precipitation has been underestimated by IPPs in wet regions and overestimated in dry regions.

e. Evaluation of the seasonal impact

Figure 11 shows the box-and-whisker plots of the IPPs and observed seasonal precipitation. Individual points in the figure refer to seasonal precipitation in a given

station. To obtain the mean seasonal precipitations, the summation of daily precipitation in a given season for each year was calculated, and then the mean is taken for that season over the duration of the study period (four values for each season in each station from 1 June 2014 to 31 May 2018). Figure 11 illustrates that the early and late runs show a better estimation of the mean precipitation than the final runs in winter (mean of winter precipitation for observed, V05E, V05L, V05F, V06E, V06L, and V06F are equal to 97, 100, 97, 109, 102, 102, and 113 mm, respectively). However, in the spring and summer seasons, the precipitation was significantly overestimated through the early and late runs. Further, in spring the precipitation variations were considerably higher in early and late runs of V06 compared to the observed precipitation, that is, the standard division of observed precipitation was 50 mm, whereas both V06E and V06L were equal to 88 mm (75% more than the observation). Unlike spring and summer, IPPs underestimate precipitation over Iran in autumn (see Fig. 11).

TABLE 8. Calculated assessment metrics for the IPPs in different climatic zones on a daily scale.

Climatic zone	Metrics	V05E	V05L	V05F	V06E	V06L	V06F
Extra arid and arid	KGE	−0.43	−0.37	0.33	−0.91	−0.95	0.28
	CC	0.58	0.60	0.71	0.61	0.63	0.68
	RBias (%)	135	129	55	186	190	63
	RMSE (mm)	3.25	3.23	1.87	4.12	4.32	2.20
	SE (%)	1	2	5	6	8	1
	RE (%)	99	98	95	94	92	99
	POD	0.64	0.65	0.62	0.64	0.64	0.60
	FAR	0.50	0.45	0.37	0.53	0.50	0.39
	CSI	0.39	0.42	0.46	0.37	0.39	0.44
	HSS	0.52	0.56	0.60	0.50	0.53	0.58
Semiarid	KGE	0.50	0.56	0.62	0.41	0.42	0.61
	CC	0.62	0.64	0.73	0.60	0.61	0.71
	RBias (%)	32	26	24	44	43	25
	RMSE (mm)	3.53	3.30	2.67	3.99	4.04	2.78
	SE (%)	4	5	6	2	1	6
	RE (%)	96	95	94	98	99	94
	POD	0.56	0.56	0.59	0.57	0.56	0.58
	FAR	0.36	0.32	0.30	0.38	0.37	0.33
	CSI	0.43	0.44	0.47	0.42	0.42	0.45
	HSS	0.52	0.54	0.57	0.51	0.51	0.55
Mediterranean	KGE	0.36	0.49	0.68	0.34	0.35	0.66
	CC	0.73	0.72	0.81	0.74	0.74	0.80
	RBias (%)	58	42	22	60	59	22
	RMSE (mm)	4.37	4.20	2.59	4.00	4.18	2.60
	SE (%)	3	1	6	3	3	9
	RE (%)	97	99	94	97	97	91
	POD	0.69	0.69	0.70	0.69	0.68	0.68
	FAR	0.35	0.31	0.27	0.37	0.37	0.31
	CSI	0.51	0.53	0.56	0.49	0.49	0.52
	HSS	0.60	0.63	0.66	0.58	0.58	0.62
Subhumid	KGE	0.65	0.66	0.66	0.62	0.60	0.68
	CC	0.69	0.68	0.71	0.63	0.61	0.70
	RBias (%)	−15	−13	−16	3	9	−12
	RMSE (mm)	5.13	5.24	4.93	6.02	6.65	5.02
	SE (%)	31	25	34	17	8	31
	RE (%)	69	75	66	83	92	69
	POD	0.47	0.49	0.51	0.51	0.51	0.52
	FAR	0.40	0.36	0.34	0.43	0.42	0.37
	CSI	0.36	0.38	0.40	0.37	0.37	0.39
	HSS	0.41	0.45	0.47	0.42	0.42	0.46
Humid	KGE	0.53	0.50	0.57	0.54	0.49	0.60
	CC	0.62	0.62	0.67	0.58	0.55	0.68
	RBias (%)	−27	−31	−27	−17	−21	−23
	RMSE (mm)	5.41	5.38	5.09	5.91	6.08	5.03
	SE (%)	44	48	42	30	32	41
	RE (%)	56	52	58	70	68	59
	POD	0.51	0.51	0.55	0.53	0.52	0.59
	FAR	0.33	0.28	0.24	0.35	0.30	0.25
	CSI	0.41	0.43	0.47	0.41	0.43	0.50
	HSS	0.49	0.52	0.57	0.49	0.51	0.59
Per-humid A and B	KGE	0.04	0.05	0.27	0.09	0.09	0.26
	CC	0.29	0.31	0.45	0.33	0.35	0.44
	RBias (%)	−64	−65	−45	−61	−62	−45
	RMSE (mm)	12.85	12.78	12.08	12.70	12.59	12.24
	SE (%)	87	86	64	83	82	60
	RE (%)	13	14	36	17	18	40
	POD	0.42	0.41	0.45	0.39	0.39	0.43
	FAR	0.33	0.27	0.27	0.33	0.28	0.28
	CSI	0.35	0.35	0.39	0.33	0.34	0.37
	HSS	0.35	0.37	0.41	0.33	0.35	0.39

FIG. 10. Frequency of precipitation less than (a) 1 mm, (b) between 1 and 5 mm, (c) between 5 and 20 mm, and (d) greater than 20 mm per station in different climatic zones for observed precipitation and the IPPs for the studied stations.

FIG. 11. Box-and-whisker plots of seasonal precipitation for observed precipitation and the IPPs for the studied stations from 1 Jun 2014 to 31 May 2018: (a) winter, (b) spring, (c) summer, and (d) autumn.

Table 9 refers to the assessment metrics for IPPs in different seasons on a daily scale. Results from Table 9 denote that all IPPs in winter, spring, and summer overestimate precipitation up to 105% (except V05F in summer), whereas underestimating in autumn is up to 24%. With respect to KGE, IPPs showed a better performance in final runs across all seasons compared to early and final runs; however, this improvement in spring and summer is more significant. The largest SEs are found in summer and autumn, whereas more than 84% of errors in winter and spring are caused by REs. The highest values of FAR are observable in summer with the number of rainy days being significantly lower than other seasons throughout Iran.

Figure 12 illustrates the spatial distribution of SEs and REs in different seasons for all IPPs on a daily scale. In most stations, in all seasons as well as both versions and all runs, more than 80% of errors are REs. However, large SE in the south of the Caspian Sea cannot be ignored, especially in the early and late runs. In winter and autumn, SEs in the south of the Caspian Sea rise above 90% in early and late runs, which improve in the final run. Among all seasons, the minimum amount of SE belongs to spring with an average of SE ranging from 17% to 23% in different IPPs. In general, based on the median of SE, in winter, SEs reduce from early runs to

final runs but increase in the spring and summer, while being almost constant in autumn.

6. Discussion

According to the aforementioned analysis, it is interesting to see that IMERG V05 performed better than V06 over Iran (see Fig. 5). The same story has happened with V03 and V04. Zhao et al. (2018) and Wei et al. (2018b) showed that V03 outperforms V04 over China. Satgé et al. (2018), who compared IPPs V03, V04, and V05 over Pakistan, indicated V04 works worse than V03. However, V05, with a real enhancement, estimates precipitation with more accuracy in comparison with V03 and V04. Sharifi et al. (2019) and Anjum et al. (2019) could not find any improvement in V06 compared to the V05 across Austria and northwestern China, respectively. However, it is reported that the retrieval algorithm of IMERG V06 has benefited from several major improvements compared to V05, that is, a new morphing approach, advancements in the calibration, and refinement in the CMORPH–Kalman filter (Anjum et al. 2019). It should be noted that, in the new morphing scheme, the MERRA-2 variables are used instead of infrared data (Huffman et al. 2015). Tan et al. (2019) showed that the new morphing scheme incorporated in

TABLE 9. Calculated assessment metrics for the IPPs in different seasons on a daily scale.

Season	Metrics	V05E	V05L	V05F	V06E	V06L	V06F
Winter	KGE	0.47	0.47	0.63	0.53	0.52	0.62
	CC	0.49	0.50	0.67	0.54	0.54	0.66
	RBias (%)	4	0	13	5	6	16
	RMSE (mm)	5.01	4.96	3.74	4.67	4.83	3.97
	SE (%)	16	15	15	14	11	10
	RE (%)	84	85	85	86	89	90
	POD	0.48	0.48	0.54	0.47	0.47	0.54
	FAR	0.39	0.35	0.33	0.39	0.37	0.35
	CSI	0.37	0.38	0.43	0.36	0.37	0.42
	HSS	0.45	0.47	0.51	0.44	0.45	0.50
Spring	KGE	0.20	0.30	0.59	−0.14	−0.12	0.57
	CC	0.59	0.62	0.70	0.57	0.58	0.68
	RBias (%)	67	58	24	105	103	25
	RMSE (mm)	4.27	4.09	2.91	6.14	6.24	3.09
	SE (%)	4	3	11	3	3	10
	RE (%)	96	97	89	97	97	90
	POD	0.64	0.65	0.64	0.65	0.65	0.63
	FAR	0.39	0.33	0.27	0.40	0.37	0.30
	CSI	0.45	0.49	0.52	0.46	0.47	0.50
	HSS	0.52	0.57	0.61	0.53	0.55	0.58
Summer	KGE	0.12	0.08	0.54	0.10	0.02	0.48
	CC	0.36	0.39	0.56	0.34	0.35	0.58
	RBias (%)	49	58	−2	47	61	1
	RMSE (mm)	2.38	2.35	1.86	2.32	2.40	1.77
	SE (%)	34	32	37	42	36	51
	RE (%)	66	68	63	58	64	49
	POD	0.49	0.49	0.43	0.46	0.47	0.43
	FAR	0.51	0.53	0.42	0.56	0.59	0.46
	CSI	0.32	0.32	0.33	0.29	0.28	0.31
	HSS	0.46	0.46	0.48	0.42	0.41	0.46
Autumn	KGE	0.39	0.39	0.48	0.40	0.42	0.48
	CC	0.45	0.46	0.54	0.44	0.44	0.52
	RBias (%)	−20	−24	−19	−7	−9	−16
	RMSE (mm)	5.93	5.81	5.50	6.09	6.18	5.66
	SE (%)	58	61	54	51	47	48
	RE (%)	42	39	46	49	53	52
	POD	0.57	0.56	0.58	0.57	0.54	0.56
	FAR	0.37	0.33	0.31	0.44	0.42	0.33
	CSI	0.43	0.44	0.46	0.39	0.39	0.44
	HSS	0.55	0.56	0.59	0.50	0.50	0.56

IMERG V06 has a better performance globally compared to V05; however, they could not find any clear improvement over the conterminous United States. The relatively poor performance of MERRA-2 simulations over Iran compared to the global average (Reichle et al. 2017) might affect the performance of the new morphing scheme over the country. Further research is needed to express a general judgment about the poorer performance of the IMERG V06 compared to its previous version. However, we expect that the IMERG team would consider the result of current study as an alert to reconsider the retrieval algorithm in the next version.

Our assessments indicated there is no significant improvement in late runs of IPPs in comparison with their early runs, while the advantage of final runs is a

considerable improvement compared to early and late runs (see Figs. 4 and 5). The same results were reported by Sharifi et al. (2018) and O et al. (2017) over the northeast and southeastern Austria, respectively. This should be noted that the Global Precipitation Climatology Centre (GPCC) dataset is used for adjusting the IMERG final run. Hence, the better performance of that is attributed to applying the gauge adjustment (Ramsauer et al. 2018; Yuan et al. 2018). Contrary to our results, Tan and Santo (2018) stated there are no significant improvements in the IMERG final run compared to near-real-time products over Malaysia. Their results may have been influenced by the improper performance of GPCC over Malaysia. In fact, they showed that the GPCC product with only

FIG. 12. Spatial distribution of systematic and random errors for the IPPs for the studied stations in different seasons on a daily scale.

24 gauges (in recent years) could not capture the spatial variability of precipitation across Malaysia.

However, although there is a significant improvement in final runs, they still overestimate the precipitation from

January to June and underestimate it between August and November (see Fig. 4b). Besides the bias, other assessment metrics like CCs still need to improve (see Fig. 5). The possible reasons for these discrepancies include the following:

- (i) The GPCC data which are used for the adjustment of final runs have a bias over Iran (Hosseini-Moghari et al. 2018), and therefore are not able to remove the bias of IPPs completely.
- (ii) Currently the GPCC Monitoring Product (version 6) is applied for adjusting IPPs following 2016 because the GPCC Full Data Reanalysis is available for the period of 1891–2016 (Huffman et al. 2015). The GPCC Monitoring Product uses less observed data compared with the GPCC Full Data Reanalysis. The GPCC Monitoring Product is developed based on 7000 stations while the GPCC Full Data Reanalysis is based on data from 67 200 stations worldwide (Schneider et al. 2014; Schneider et al. 2017). Therefore, the performance of IPPs following 2016 (more than half of the time period of this study) could be worse than before 2016.
- (iii) The IPPs are adjusted at a monthly scale (Huffman et al. 2015), while in this study the analyses are carried out based on a daily scale. So, improvement in the performance of IPP final runs on a daily scale would be less than a monthly scale.
- (iv) GPCC Full Data Reanalysis and GPCC Monitoring Product have spatial resolutions of $0.5^\circ \times 0.5^\circ$ and $1^\circ \times 1^\circ$, respectively, while the IPPs provide the precipitation data at $0.1^\circ \times 0.1^\circ$ spatial resolution. Therefore, some errors can arise from this difference in spatial resolution.

With respect to Fig. 5, with a median of CCs equal to 0.6 (for early and late runs) and 0.7 (for final), we concluded that there is generally a moderate correlation [$0.5 < CC < 0.7$ was considered as a moderate correlation based on Tan and Santo (2018)] between IPPs and gauge observations. The same results have been reported for other regions, for example, Blue Nile basin (Sahlu et al. 2016), Mekong River basin (Wang et al. 2017), Malaysia (Tan and Santo 2018), and northern Pakistan (Anjum et al. 2018). Therefore, following the Brown (2006) and Condom et al. (2011) who suggested a correlation higher than 0.7 and RBias between $\pm 10\%$ for a reliable precipitation dataset, none of the IPPs (without preprocessing) can be considered as an alternative for gauge observations over Iran (see Table 7). Also, based on Fig. 5, we showed the gauge adjustment could not modify the POD, while FAR, CSI, and HSS improved in final runs, which is in line with Gosset et al. (2013) and Behrangi et al. (2014a). It is worth mentioning that the gauge adjustment amends the amounts of precipitation, not the occurrence of that (Gosset et al. 2013). Therefore, it is possible that the precipitation amount is sited to zero after the gauge adjustment to modify the precipitation volume. In this case, the

number of falsely detected events would be reduced. As a result, FAR, CSI, and HSS (that depend on the falsely detected events, unlike POD) can be improved after the gauge adjustment (final runs). We also analyzed the sensitivity of categorical validation metrics to selected rain/no-rain threshold. Two thresholds of 0.1 and 0.5 mm day^{-1} were considered for the IPPs. The results illustrated that the categorical validation metrics highly depend on the selection of rainfall intensity threshold. Based on two new thresholds both POD and FAR increased. For example, by considering the thresholds of 0.1, 0.5, and 1 mm day^{-1} , the median of POD and FAR for IMERG V06F have been 0.80, 0.67, and 0.60 (for POD) and 0.56, 0.40, and 0.32 (for FAR), respectively.

Our evaluations revealed a weak dependence of IPPs to the elevation, particularly with respect to categorical validation metrics (see Fig. 8 and Table 6). It should be noted that, in this study, the number of stations in different elevation classes is not equal, which may cause uncertainties to our evaluations. However, according to Fig. 8, without considering the elevation classes, there is no significant relationship between elevation and assessment metrics. This result is not entirely consistent with previous researches, but it is not in conflict with them either. Xu et al. (2017), who assessed the effects of elevation on accumulative rainfall over southern Tibetan Plateau, indicated unlike TRMM, there is no clear evidence for the influence of topographic on rainfall retrieval algorithm of GPM (IMERG). However, Beria et al. (2017) reported an inverse relationship between IMERG skill and the elevation over India. Fang et al. (2019) also found this relationship, but generally weak, over China. Here, we point out some factors that might be generally changed by altitude but not necessarily, which may not be noticed:

- (i) Precipitation is generally increased with an increase of elevation. Therefore, a more significant relationship between the intensity of precipitation and performance of IPPs can exist.
- (ii) The number of gauges is scarce in high mountainous area, therefore after gauge adjustment, the performance of IPPs may not be significantly increased in this area (Chen and Li 2016).
- (iii) Precipitation processes are usually complex in high-elevation regions. Therefore, its estimation will be more complicated for satellite sensors (Fang et al. 2019).
- (iv) When the surface is covered by snow or ice (which is common in high mountainous regions), due to some difficulty in retrievals algorithms, IMERG uses only infrared-based estimations that are much less certain than the microwave estimations (Zhou et al. 2019).

Our results indicated (see Fig. 9) that IPPs tend to underestimate the frequency of tiny precipitation events ($<1 \text{ mm day}^{-1}$) while overestimating the frequency of light, moderate, and low heavy precipitation events ($1\text{--}20 \text{ mm day}^{-1}$). The same results also reported by Xu et al. (2017) and Tan and Santo (2018) over the Tibetan Plateau and Malaysia, respectively. Our results showed a direct relationship between SE and precipitation rate where SE is increased along with the increase of rain intensity (from 6% for precipitation $< 1 \text{ mm day}^{-1}$ to 61% for precipitation $\geq 20 \text{ mm day}^{-1}$), which is in agreement with findings of Habib et al. (2009) and AghaKouchak et al. (2012). Moreover, Table 7 shows an increase in the ability of IPPs to detect precipitation events with increasing precipitation intensity (POD ranging from 0.75 to 0.83 for precipitation $\geq 5 \text{ mm day}^{-1}$). In line with our results, Habib et al. (2009) and Kirstetter et al. (2015) stated POD could be increased with increasing precipitation rate.

Generally, IPPs overestimated the precipitation over Iran (see Figs. 4b, 5, and 6 and Table 7). Xu et al. (2017), Anjum et al. (2018), Sunilkumar et al. (2019), and Islam (2018) reported similar results over southern Tibetan Plateau, northern Pakistan, Japan and Nepal, and Bangladesh, respectively. This overestimation is higher for early and late runs compared to the final runs, especially for V06. However, the use of GPCC for adjusting the data in final runs could reduce the RBias from 17% in EV05 to 5% in FV05 and from 32% in EV06 to 7% in FV06 (see Table 7). The largest overestimation can be seen from February to May (see Fig. 4b). These results are in a good agreement with Chen and Li (2016) and O et al. (2017) who showed that the IMERG RBias varies in different months and is larger for real-time runs compared to the final runs over China and Austria, respectively. Besides general overestimation, IPPs have been underestimated precipitation between August and November. This underestimation is slightly in August but is significant between September and November, that is, autumn (see Table 9). The source of this underestimation can be found in an underestimation in per-humid and humid regions (see Table 8), where those receive a part significant of their annual precipitation in the autumn.

The performance of IPPs over Iran is quite diverse in different climates (see Fig. 10 and Table 8), that is, there is an overestimation in dry regions and underestimation in wet areas. These overestimation and underestimation mostly attributed to estimation of a large number of precipitation events higher than 5 mm day^{-1} by IPPs in the dry areas and vice versa in wet regions (see Fig. 10). The evaporation of small particles of precipitation in the lower precipitating system drives IMERG to overestimate

the event rate (Zhou et al. 2019). This eventually leads to the overestimation of IPPs in arid regions. Also, Anjum et al. (2018) argued that although the higher atmospheric concentration increases the size of cloud drops, it may also intercept the precipitation and cause the overestimation of precipitation by SPPs in northern Pakistan. However, although the concentration of aerosols in the dry regions of Iran has been increased (Arkian and Nicholson 2018; Rashki et al. 2012), many other factors should be considered to claim about their effects on decreasing precipitation (Khain 2009).

Underestimations can be seen in the south of the Caspian Sea (See Fig. 6), where almost all stations are located in humid regions (see Fig. 1). Not only IPPs, but also other SPPs have been underestimated the precipitation in this area. The same results were reported for TRMM (Darand et al. 2017; Javanmard et al. 2010), PERSIANN (Alijanian et al. 2017; Katiraei-Boroujerdy et al. 2013), and CMORPH (Alijanian et al. 2017). The leading cause of this underestimation can be found in the region's rainfall regime. Due to the proximity to the Alborz mountain chain the primary rainfall regime in the south of the Caspian Sea is orographic. The area with orographically driven events is known as a problematic area for SPPs because in this area warm rain enhancement is important, and the retrievals underestimate rates under this circumstance (Zhou et al. 2019). The limited ability of SPPs in detection of orographic precipitation is well documented in several studies (Behrangi et al. 2014b; Mehran and AghaKouchak 2014; Shige et al. 2013; Sorooshian et al. 2011). Actually, in the case of warm orographic rainfall, infrared sensors have limited ability for detecting rainfall when the cloud tops are warmer than the temperature thresholds of the sensors. And microwave sensors estimate rainfall rate based on the scattering by ice particles, which can noticeably cause underestimation (Rafiuddin et al. 2010). Over the south of the Caspian Sea, no significant improvements of IPPs in the final run can be associated with an unacceptable performance of GPCC dataset for this region (Hosseini-Moghari et al. 2018) and resulted in an SE higher than 60% (see Table 8).

Assessing the impacts of seasonal changes revealed that the worst results were related to the summer season followed by the autumn (see Table 9). Higher FAR values in the summer represent the higher frequency of false rainfall in dry seasons. The low performance in the summer (warm and dry season) can be attributed to the adoption of retrieval algorithms for capturing warm rainfall and ignoring icy cirrus clouds that have no rain (Behrangi et al. 2010, 2009). Generally, the winter and spring are considered rainy seasons over most of Iran, while the south of the Caspian Sea receives a significant

portion of its annual rainfall in the autumn. Therefore, negative RBias values in the autumn (see Table 9) can be linked to the significant underestimation of precipitation by IPPs in this area resulted in a huge SE (see Fig. 12). A high amount of SE in the summer, autumn (see Table 9), wet regions (see Table 8), and for the heavy precipitation intensity (see Table 7) indicated that a bias removal algorithm is necessary to improve the accuracy of IPPs not only for adjusting the volume of precipitation but also for estimating the precipitation peaks.

7. Conclusions

This study attempted to shed some light on the validity of different runs (early, late, and final) and versions (V05 and V06) of IPPs over Iran. We assessed the performance of IPPs against 76 key synoptic stations on a daily scale from 1 June 2014 to 30 June 2018. Besides general evaluations, the impacts of elevation, precipitation intensity, climate, and seasonal changes were considered in the assessment. The main conclusions of this study can be summarized as follows:

- IMERG V05 performs better than V06, particularly in early and late runs. The difference between early and late runs of IPPs was negligible over Iran, however, their performance was significantly increased in the final run. For instance, the median of RBias for V05E, V05L, V05F, V06E, V06L, and V06F were equal to 52%, 45%, 25%, 62%, 60%, and 30%, respectively. Also, there was no difference between the POD value in different IPPs (the median of POD equals 0.60), while an improvement in FAR, CSI, and HSS was achieved. This suggests that the ability of detection of precipitation events could not be improved, but the number of false alarms could be decreased.
- Our assessment revealed that the dependence of IPPs to elevation is low, especially with respect to categorical validation metrics. However, the performance of IPPs strongly dependent on climate and precipitation intensity. All IPPs, especially early and late runs, tend to underestimate the frequency of “no\tiny” precipitation class ($<1 \text{ mm day}^{-1}$) but overestimate the frequency from light to low heavy rainfall (between 1 and 20 mm day^{-1}). Also, all IPPs are inclined to overestimate the amount of precipitation for classes less than 5 mm day^{-1} , and on the contrary, attempt to underestimate the amount of precipitation for classes higher than 5 mm day^{-1} (especially for more than 20 mm day^{-1}). Acceptable performance of IPPs in estimation of the frequency of high heavy rainfall reveals the capability of the products to capture the high heavy rainfall. In contrast, the large negative RBias indicates that the performance of

the IPPs in estimating the rainfall amount needs to be enhanced.

- Assessing the impacts of climatic zones illustrated that the validity and error characteristic of IPPs strongly depend on the climate. IPPs tend to overestimate and underestimate the precipitation in dry and wet areas, respectively, and resulted in less overall bias. However, IPPs still overestimate precipitation due to the greater extent of the arid regions in Iran. Also, investigation of the joint effects of the climate and precipitation intensity showed that IPPs overestimate the number of precipitation events between 1 and 5 mm day^{-1} in all climatic zones. However, the frequency of precipitation higher than 5 mm day^{-1} fluctuates between an overestimation (from the extra arid to Mediterranean climate) and an underestimation (from the subhumid to per-humid B). The overestimation (underestimation) of the moderate and heavy frequencies in dry regions (humid regions) is responsible for the overestimation (underestimation) of the total precipitation volume in these regions.
- Generally, all IPPs underestimate the amount of precipitation in the autumn while the opposite is true for other seasons. With respect to KGE and RBias, IPPs have a better performance in the winter in comparing with other seasons. In addition, it is interesting that, early and late runs in both versions have a better performance in comparison with the estimations of final runs for the total amount of precipitation in the winter which resulted in a small RBias ranging from 0% to 6% for early and late runs. However, their related RMSE is worse than the final runs so that more errors in individual events for early and late runs are suggested in comparison with the final ones.

We hope that the multifaceted analysis performed in this study would provide useful information to users and the IMERG developer team for a better understanding of the error characteristics of two available versions of IMERG products. However, one limitation of our analysis was that we only had one synoptic station inside each IMERG pixel that does not represent the whole pixel. On the other hand, according to Islam (2018) with consideration of a long time period (more than 3 years), the errors associated with the spatial representativeness of the rain gauge data are minimized. Furthermore, we believe that given the high spatial resolution of IPPs, the adopted method will be sufficient. Therefore, we assume the results of this study cannot be affected by the uncertainty caused by using the point-to-pixel approach. As a final word, notwithstanding the significant improvements that have been made in precipitation estimation based on satellite observations,

according to our results further progress on the precipitation retrieval algorithm and satellite sensors is needed by considering climatic zones and precipitation intensity influences. Therefore, assessing the validity of SPPs in different regions will still remain a work in progress.

Acknowledgments. This research was funded by the National Natural Science Foundation of China (41790424), International Partnership Program of Chinese Academy of Sciences (131A11KYSB20170113), Chinese Academy of Sciences President's International Fellowship Initiative (2019VEA0019), and Newton Advanced Fellowships. The authors are grateful to the Islamic Republic of Iran Meteorological Organization (IRIMO) and the National Aeronautics and Space Administration (NASA) for providing in situ data and GPM-IMERG products, respectively. The authors declare that they have no conflict of interests.

REFERENCES

- AghaKouchak, A., A. Mehran, H. Norouzi, and A. Behrangi, 2012: Systematic and random error components in satellite precipitation data sets. *Geophys. Res. Lett.*, **39**, L09406, <https://doi.org/10.1029/2012GL051592>.
- Alijani, B., M. Ghohroudi, and N. Arabi, 2008: Developing a climate model for Iran using GIS. *Theor. Appl. Climatol.*, **92**, 103–112, <https://doi.org/10.1007/s00704-006-0292-y>.
- Alijanian, M., G. R. Rakhshandehroo, A. K. Mishra, and M. Dehghani, 2017: Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. *Int. J. Climatol.*, **37**, 4896–4914, <https://doi.org/10.1002/joc.5131>.
- Anagnostopoulos, G., D. Koutsoyiannis, A. Christofides, A. Efstratiadis, and N. Mamassis, 2010: A comparison of local and aggregated climate model outputs with observed data. *Hydrol. Sci. J.*, **55**, 1094–1110, <https://doi.org/10.1080/02626667.2010.513518>.
- Anjum, M. N., and Coauthors, 2018: Performance evaluation of latest integrated multi-satellite retrievals for Global Precipitation Measurement (IMERG) over the northern highlands of Pakistan. *Atmos. Res.*, **205**, 134–146, <https://doi.org/10.1016/j.atmosres.2018.02.010>.
- , and Coauthors, 2019: Assessment of IMERG-V06 precipitation product over different hydro-climatic regimes in the tianshan mountains, north-western China. *Remote Sens.*, **11**, 2314, <https://doi.org/10.3390/rs11192314>.
- Arkian, F., and S. Nicholson, 2018: Long-term variations of aerosol optical depth and aerosol radiative forcing over Iran based on satellite and AERONET data. *Environ. Monit. Assess.*, **190**, 1–15, <https://doi.org/10.1007/s10661-017-6336-1>.
- Ashouri, H., K. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat, 2015: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Amer. Meteor. Soc.*, **96**, 69–83, <https://doi.org/10.1175/BAMS-D-13-00068.1>.
- Aslami, F., A. Ghorbani, B. Sobhani, and A. Esmali, 2019: Comprehensive comparison of daily IMERG and GSMaP satellite precipitation products in Ardabil Province, Iran. *Int. J. Remote Sens.*, **40**, 3139–3153, <https://doi.org/10.1080/01431161.2018.1539274>.
- Behrangi, A., K. Hsu, B. Imam, S. Sorooshian, G. J. Huffman, and R. J. Kuligowski, 2009: PERSIANN-MSA: A precipitation estimation method from satellite-based multispectral analysis. *J. Hydrometeor.*, **10**, 1414–1429, <https://doi.org/10.1175/2009JHM1139.1>.
- , —, —, and —, 2010: Daytime precipitation estimation using bispectral cloud classification system. *J. Appl. Meteor. Climatol.*, **49**, 1015–1031, <https://doi.org/10.1175/2009JAMC2291.1>.
- , Y. Tian, B. H. Lambriksen, and G. L. Stephens, 2014a: What does CloudSat reveal about global land precipitation detection by other spaceborne sensors? *Water Resour. Res.*, **50**, 4893–4905, <https://doi.org/10.1002/2013WR014566>.
- , K. Andreadis, J. B. Fisher, F. J. Turk, S. Granger, T. Painter, and N. Das, 2014b: Satellite-based precipitation estimation and its application for streamflow prediction over mountainous western U.S. basins. *J. Appl. Meteor. Climatol.*, **53**, 2823–2842, <https://doi.org/10.1175/JAMC-D-14-0056.1>.
- Beria, H., T. Nanda, D. S. Bisht, and C. Chatterjee, 2017: Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale. *Hydrol. Earth Syst. Sci.*, **21**, 6117–6134, <https://doi.org/10.5194/hess-21-6117-2017>.
- Brown, J. E., 2006: An analysis of the performance of hybrid infrared and microwave satellite precipitation algorithms over India and adjacent regions. *Remote Sens. Environ.*, **101**, 63–81, <https://doi.org/10.1016/j.rse.2005.12.005>.
- Camici, S., L. Ciabatta, C. Massari, and L. Brocca, 2018: How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean area. *J. Hydrol.*, **563**, 950–961, <https://doi.org/10.1016/j.jhydrol.2018.06.067>.
- Chen, F., and X. Li, 2016: Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.*, **8**, 472, <https://doi.org/10.3390/rs8060472>.
- Condom, T., P. Rau, and J. C. Espinoza, 2011: Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007. *Hydrol. Processes*, **25**, 1924–1933, <https://doi.org/10.1002/hyp.7949>.
- Darand, M., J. Amanollahi, and S. Zandkarimi, 2017: Evaluation of the performance of TRMM Multi-satellite Precipitation Analysis (TMPA) estimation over Iran. *Atmos. Res.*, **190**, 121–127, <https://doi.org/10.1016/j.atmosres.2017.02.011>.
- De Martonne, E., 1926: Une Nouvelle fonction climatologique: L'Indice d'aridité. *Meteorologie*, **2**, 449–458.
- Dezfooli, D., B. Abdollahi, S.-M. Hosseini-Moghari, and K. Ebrahimi, 2018: A comparison between high-resolution satellite precipitation estimates and gauge measured data: Case study of Gorganrood basin, Iran. *J. Water Supply Res. Technol.*, **67**, 236–251, <https://doi.org/10.2166/aqua.2018.062>.
- Diem, J. E., J. Hartter, S. J. Ryan, and M. W. Palace, 2014: Validation of satellite rainfall products for western Uganda. *J. Hydrometeor.*, **15**, 2030–2038, <https://doi.org/10.1175/JHM-D-13-0193.1>.
- Duan, Z., J. Liu, Y. Tuo, G. Chiogna, and M. Disse, 2016: Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Sci. Total Environ.*, **573**, 1536–1553, <https://doi.org/10.1016/j.scitotenv.2016.08.213>.
- Fang, J., W. Yang, Y. Luan, J. Du, A. Lin, and L. Zhao, 2019: Evaluation of the TRMM 3B42 and GPM IMERG

- products for extreme precipitation analysis over China. *Atmos. Res.*, **223**, 24–38, <https://doi.org/10.1016/j.atmosres.2019.03.001>.
- Gehne, M., T. M. Hamill, G. N. Kiladis, and K. E. Trenberth, 2016: Comparison of global precipitation estimates across a range of temporal and spatial scales. *J. Climate*, **29**, 7773–7795, <https://doi.org/10.1175/JCLI-D-15-0618.1>.
- Ghajarnia, N., A. Liaghat, and P. D. Arasteh, 2015: Comparison and evaluation of high resolution precipitation estimation products in Urmia Basin-Iran. *Atmos. Res.*, **158–159**, 50–65, <https://doi.org/10.1016/j.atmosres.2015.02.010>.
- Gosset, M., J. Viarre, G. Quantin, and M. Alcobá, 2013: Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Quart. J. Roy. Meteor. Soc.*, **139**, 923–940, <https://doi.org/10.1002/qj.2130>.
- Guo, H., S. Chen, A. Bao, J. Hu, B. Yang, and P. Stepanian, 2016: Comprehensive evaluation of high-resolution satellite-based precipitation products over China. *Atmosphere*, **7**, 6, <https://doi.org/10.3390/atmos7010006>.
- Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez, 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>.
- Habib, E., A. Henschke, and R. F. Adler, 2009: Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmos. Res.*, **94**, 373–388, <https://doi.org/10.1016/j.atmosres.2009.06.015>.
- Hazra, A., V. Maggioni, P. Houser, H. Antil, and M. Noonan, 2019: A Monte Carlo-based multi-objective optimization approach to merge different precipitation estimates for land surface modeling. *J. Hydrol.*, **570**, 454–462, <https://doi.org/10.1016/j.jhydrol.2018.12.039>.
- Hosseini-Moghari, S.-M., S. Araghinejad, and K. Ebrahimi, 2018: Spatio-temporal evaluation of global gridded precipitation datasets across Iran. *Hydrol. Sci. J.*, **63**, 1669–1688, <https://doi.org/10.1080/02626667.2018.1524986>.
- Hou, A. Y., and Coauthors, 2014: The global precipitation measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Huffman, G. J., 1997: Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteor.*, **36**, 1191–1201, [https://doi.org/10.1175/1520-0450\(1997\)036<1191:EORMSR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<1191:EORMSR>2.0.CO;2).
- , and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , D. T. Bolvin, and E. J. Nelkin, 2015: Integrated Multi-satellite Retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code 612 Tech. Doc., 48 pp., http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf.
- Islam, M. A., 2018: Statistical comparison of satellite-retrieved precipitation products with rain gauge observations over Bangladesh. *Int. J. Remote Sens.*, **39**, 2906–2936, <https://doi.org/10.1080/01431161.2018.1433890>.
- Javanmard, S., A. Yatagai, M. I. Nodzu, J. BodaghJamali, and H. Kawamoto, 2010: Comparing high-resolution gridded precipitation data with satellite rainfall estimates of TRMM_3B42 over Iran. *Adv. Geosci.*, **25**, 119–125, <https://doi.org/10.5194/adgeo-25-119-2010>.
- Joyce, R. J., 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, [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Katiraei-Boroujerdy, P.-S., N. Nasrollahi, K. Hsu, and S. Sorooshian, 2013: Evaluation of satellite-based precipitation estimation over Iran. *J. Arid Environ.*, **97**, 205–219, <https://doi.org/10.1016/j.jaridenv.2013.05.013>.
- Khain, A., 2009: Notes on state-of-the-art investigations of aerosol effects on precipitation: A critical review. *Environ. Res. Lett.*, **4**, 015004, <https://doi.org/10.1088/1748-9326/4/1/015004>.
- Khalili, A., and J. Rahimi, 2014: High-resolution spatiotemporal distribution of precipitation in Iran: A comparative study with three global-precipitation datasets. *Theor. Appl. Climatol.*, **118**, 211–221, <https://doi.org/10.1007/s00704-013-1055-1>.
- , and —, 2018: Climate. *The Soils of Iran*, M. Roozitalab, H. Siadat, and A. Farshad, Eds., World Soils Book Series, Springer, 19–33, https://doi.org/10.1007/978-3-319-69048-3_3.
- Khodadoust Siuki, S., B. Saghafian, and S. Moazami, 2017: Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products. *Int. J. Remote Sens.*, **38**, 558–571, <https://doi.org/10.1080/01431161.2016.1268735>.
- Kidd, C., and G. Huffman, 2011: Global precipitation measurement. *Meteor. Appl.*, **18**, 334–353, <https://doi.org/10.1002/met.284>.
- Kirstetter, P. E., Y. Hong, J. Gourley, M. Schwaller, W. Petersen, and Q. Cao, 2015: Impact of sub-pixel rainfall variability on spaceborne precipitation estimation: Evaluating the TRMM 2A25 product. *Quart. J. Roy. Meteor. Soc.*, **141**, 953–966, <https://doi.org/10.1002/qj.2416>.
- Kling, H., M. Fuchs, and M. Paulin, 2012: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.*, **424–425**, 264–277, <https://doi.org/10.1016/j.jhydrol.2012.01.011>.
- Lai, C., R. Zhong, Z. Wang, X. Wu, X. Chen, P. Wang, and Y. Lian, 2019: Monitoring hydrological drought using long-term satellite-based precipitation data. *Sci. Total Environ.*, **649**, 1198–1208, <https://doi.org/10.1016/j.scitotenv.2018.08.245>.
- Lee, J., E.-H. Lee, and K.-H. Seol, 2019: Validation of Integrated Multisatellite Retrievals for GPM (IMERG) by using gauge-based analysis products of daily precipitation over East Asia. *Theor. Appl. Climatol.*, **137**, 2497–2512, <https://doi.org/10.1007/s00704-018-2749-1>.
- Ma, Y., Y. Zhang, D. Yang, and S. B. Farhan, 2015: Precipitation bias variability versus various gauges under different climatic conditions over the Third Pole Environment (TPE) region. *Int. J. Climatol.*, **35**, 1201–1211, <https://doi.org/10.1002/joc.4045>.
- Maggioni, V., and C. Massari, 2018: On the performance of satellite precipitation products in riverine flood modeling: A review. *J. Hydrol.*, **558**, 214–224, <https://doi.org/10.1016/j.jhydrol.2018.01.039>.
- , M. R. Sapiiano, and R. F. Adler, 2016: Estimating uncertainties in high-resolution satellite precipitation products: Systematic or random error? *J. Hydrometeorol.*, **17**, 1119–1129, <https://doi.org/10.1175/JHM-D-15-0094.1>.
- Mehran, A., and A. AghaKouchak, 2014: Capabilities of satellite precipitation datasets to estimate heavy precipitation rates at different temporal accumulations. *Hydrol. Processes*, **28**, 2262–2270, <https://doi.org/10.1002/hyp.9779>.
- O, S., U. Foelsche, G. Kirchengast, J. Fuchsberger, J. Tan, and W. A. Petersen, 2017: Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge

- data in southeastern Austria. *Hydrol. Earth Syst. Sci.*, **21**, 6559–6572, <https://doi.org/10.5194/hess-21-6559-2017>.
- Persson, A., 2001: User guide to ECMWF forecast products. Meteorological Bulletin M3.2, ECMWF, 115 pp.
- Prakash, S., 2019: Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *J. Hydrol.*, **571**, 50–59, <https://doi.org/10.1016/j.jhydrol.2019.01.036>.
- , A. K. Mitra, A. AghaKouchak, Z. Liu, H. Norouzi, and D. Pai, 2018: A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.*, **556**, 865–876, <https://doi.org/10.1016/j.jhydrol.2016.01.029>.
- Rafiuddin, M., H. Uyeda, and M. N. Islam, 2010: Characteristics of monsoon precipitation systems in and around Bangladesh. *Int. J. Climatol.*, **30**, 1042–1055, <https://doi.org/10.1002/joc.1949>.
- Rahimi, J., M. Ebrahimpour, and A. Khalili, 2013: Spatial changes of extended De Martonne climatic zones affected by climate change in Iran. *Theor. Appl. Climatol.*, **112**, 409–418, <https://doi.org/10.1007/s00704-012-0741-8>.
- Ramsauer, T., T. Weiß, and P. Marzahn, 2018: Comparison of the GPM IMERG final precipitation product to RADOLAN weather radar data over the topographically and climatically diverse Germany. *Remote Sens.*, **10**, 2029, <https://doi.org/10.3390/rs10122029>.
- Rashki, A., D. G. Kaskaoutis, C. Rautenbach, P. G. Eriksson, M. Qiang, and P. Gupta, 2012: Dust storms and their horizontal dust loading in the Sistan region, Iran. *Aeolian Res.*, **5**, 51–62, <https://doi.org/10.1016/j.aeolia.2011.12.001>.
- Reichle, R. H., C. S. Draper, Q. Liu, M. Girotto, S. P. Mahanama, R. D. Koster, and G. J. De Lannoy, 2017: Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, **30**, 2937–2960, <https://doi.org/10.1175/JCLI-D-16-0720.1>.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's modern-era retrospective analysis for research and applications. *J. Climate*, **24**, 3624–3648, <https://doi.org/10.1175/JCLI-D-11-00015.1>.
- Rivera, J. A., G. Marianetti, and S. Hinrichs, 2018: Validation of CHIRPS precipitation dataset along the Central Andes of Argentina. *Atmos. Res.*, **213**, 437–449, <https://doi.org/10.1016/j.atmosres.2018.06.023>.
- Rozante, J., D. Vila, J. Barboza Chiquetto, A. Fernandes, and D. Souza Alvim, 2018: Evaluation of TRMM/GPM blended daily products over Brazil. *Remote Sens.*, **10**, 882, <https://doi.org/10.3390/rs10060882>.
- Sahlu, D., E. I. Nikolopoulos, S. A. Moges, E. N. Anagnostou, and D. Hailu, 2016: First evaluation of the day-1 IMERG over the upper Blue Nile Basin. *J. Hydrometeorol.*, **17**, 2875–2882, <https://doi.org/10.1175/JHM-D-15-0230.1>.
- Satgé, F., Y. Hussain, M.-P. Bonnet, B. Hussain, H. Martinez-Carvajal, G. Akhter, and R. Uagoda, 2018: Benefits of the successive GPM based satellite precipitation estimates IMERG-V03,–V04,–V05 and GSMaP-V06,–V07 over diverse geomorphic and meteorological regions of Pakistan. *Remote Sens.*, **10**, 1373, <https://doi.org/10.3390/rs10091373>.
- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2014: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.*, **115**, 15–40, <https://doi.org/10.1007/s00704-013-0860-x>.
- , P. Finger, A. Meyer-Christoffer, E. Rustemeier, M. Ziese, and A. Becker, 2017: Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, **8**, 52, <https://doi.org/10.3390/atmos8030052>.
- Sharifi, E., R. Steinacker, and B. Saghafian, 2016: Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sens.*, **8**, 135, <https://doi.org/10.3390/rs8020135>.
- , —, and —, 2018: Multi time-scale evaluation of high-resolution satellite-based precipitation products over north-east of Austria. *Atmos. Res.*, **206**, 46–63, <https://doi.org/10.1016/j.atmosres.2018.02.020>.
- , J. Eitzinger, and W. Dorigo, 2019: Performance of the state-of-the-art gridded precipitation products over mountainous terrain: A regional study over Austria. *Remote Sens.*, **11**, 2018, <https://doi.org/10.3390/rs11172018>.
- Shige, S., S. Kida, H. Ashiwake, T. Kubota, and K. Aonashi, 2013: Improvement of TMI rain retrievals in mountainous areas. *J. Appl. Meteor. Climatol.*, **52**, 242–254, <https://doi.org/10.1175/JAMC-D-12-074.1>.
- Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, [https://doi.org/10.1175/1520-0477\(2000\)081<2035:EOPSE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSE>2.3.CO;2).
- , and Coauthors, 2011: Advanced concepts on remote sensing of precipitation at multiple scales. *Bull. Amer. Meteor. Soc.*, **92**, 1353–1357, <https://doi.org/10.1175/2011BAMS3158.1>.
- Sunilkumar, K., A. Yatagai, and M. Masuda, 2019: Preliminary Evaluation of GPM-IMERG rainfall estimates over three distinct climate zones with APHRODITE. *Earth Space Sci.*, **6**, 1321–1335, <https://doi.org/10.1029/2018EA000503>.
- Tabari, H., and P. H. Talaee, 2011: Temporal variability of precipitation over Iran: 1966–2005. *J. Hydrol.*, **396**, 313–320, <https://doi.org/10.1016/j.jhydrol.2010.11.034>.
- Tan, J., G. J. Huffman, D. T. Bolvin, and E. J. Nelkin, 2019: IMERG V06: Changes to the morphing algorithm. *J. Atmos. Oceanic Technol.*, **36**, 2471–2482, <https://doi.org/10.1175/JTECH-D-19-0114.1>.
- Tan, M. L., and H. Santo, 2018: Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia. *Atmos. Res.*, **202**, 63–76, <https://doi.org/10.1016/j.atmosres.2017.11.006>.
- , A. Ibrahim, Z. Duan, A. Cracknell, and V. Chaplot, 2015: Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sens.*, **7**, 1504–1528, <https://doi.org/10.3390/rs70201504>.
- Wang, F., H. Yang, Z. Wang, Z. Zhang, and Z. Li, 2019: Drought evaluation with CMORPH satellite precipitation data in the Yellow River Basin by using gridded standardized precipitation evapotranspiration index. *Remote Sens.*, **11**, 485, <https://doi.org/10.3390/rs11050485>.
- Wang, W., H. Lu, T. Zhao, L. Jiang, and J. Shi, 2017: Evaluation and comparison of daily rainfall from latest GPM and TRMM products over the Mekong River Basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **10**, 2540–2549, <https://doi.org/10.1109/JSTARS.2017.2672786>.
- Wei, G., H. Lü, W. T. Crow, Y. Zhu, J. Wang, and J. Su, 2018a: Comprehensive evaluation of GPM-IMERG, CMORPH, and TMPA precipitation products with gauged rainfall over Mainland China. *Adv. Meteor.*, **2018**, 1–18, <https://doi.org/10.1155/2018/3024190>.
- , —, —, —, —, and —, 2018b: Evaluation of satellite-based precipitation products from IMERG V04A

- and V03D, CMORPH and TMPA with gauged rainfall in three climatologic zones in China. *Remote Sens.*, **10**, 30, <https://doi.org/10.3390/rs10010030>.
- Willmott, C. J., 1981: On the validation of models. *Phys. Geogr.*, **2**, 184–194, <https://doi.org/10.1080/02723646.1981.10642213>.
- WMO, 2009: Handbook on CLIMAT and CLIMAT TEMP reporting. WWW Tech. Rep. WMO/TD-1188, 115 pp., https://www.wmo.int/pages/prog/www/OSY/Publications/TD1188/HandbookCLIMAT-CLIMATTEMP_en.pdf.
- Wolters, E. L. A., B. J. J. M. Van Den Hurk, and R. A. Roebeling, 2011: Evaluation of rainfall retrievals from SEVIRI reflectances over West Africa using TRMM-PR and CMORPH. *Hydrol. Earth Syst. Sci.*, **15**, 437–451, <https://doi.org/10.5194/hess-15-437-2011>.
- Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.*, **122**, 910–924, <https://doi.org/10.1002/2016JD025418>.
- Yang, X., B. Yong, Y. Hong, S. Chen, and X. Zhang, 2016: Error analysis of multi-satellite precipitation estimates with an independent raingauge observation network over a medium-sized humid basin. *Hydrol. Sci. J.*, **61**, 1813–1830, <https://doi.org/10.1080/02626667.2015.1040020>.
- Yazdanpanah, H., J. Eitzinger, and M. Baldi, 2017: Analysis of the extreme heat events in Iran. *Int. J. Climate Change Str. Manage.*, **9**, 418–432, <https://doi.org/10.1108/IJCCSM-04-2016-0046>.
- Yong, B., J. Wang, L. Ren, Y. You, P. Xie, and Y. Hong, 2016: Evaluating four multisatellite precipitation estimates over the Diaoyu Islands during Typhoon seasons. *J. Hydrometeor.*, **17**, 1623–1641, <https://doi.org/10.1175/JHM-D-15-0165.1>.
- Yuan, F., and Coauthors, 2018: Evaluation of hydrological utility of IMERG Final run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. *J. Hydrol.*, **567**, 696–711, <https://doi.org/10.1016/j.jhydrol.2018.06.045>.
- Zappa, M., 2008: Objective quantitative spatial verification of distributed snow cover simulations—An experiment for the whole of Switzerland/Vérification quantitative spatiale objective de simulations distribuées de la couche de neige—Une étude pour l'ensemble de la Suisse. *Hydrol. Sci. J.*, **53**, 179–191, <https://doi.org/10.1623/hysj.53.1.179>.
- Zhang, X., Z. Su, J. Lv, W. Liu, M. Ma, J. Peng, and G. Leng, 2019: A set of satellite-based near real-time meteorological drought monitoring data over China. *Remote Sens.*, **11**, 453, <https://doi.org/10.3390/rs11040453>.
- Zhao, H., S. Yang, S. You, Y. Huang, Q. Wang, and Q. Zhou, 2018: Comprehensive evaluation of two successive V3 and V4 IMERG final run precipitation products over Mainland China. *Remote Sens.*, **10**, 34, <https://doi.org/10.3390/rs10010034>.
- Zhou, Y., K. Nelson, K. I. Mohr, G. J. Huffman, R. Levy, and M. Grecu, 2019: A spatial-temporal extreme precipitation database from GPM IMERG. *J. Geophys. Res. Atmos.*, **124**, 10 344–10 363, <https://doi.org/10.1029/2019JD030449>.
- Zhu, Q., Y. Luo, D. Zhou, Y.-P. Xu, G. Wang, and H. Gao, 2019: Drought monitoring utility using satellite-based precipitation products over the Xiang River Basin in China. *Remote Sens.*, **11**, 1483, <https://doi.org/10.3390/rs11121483>.
- Zorzetto, E., and M. Marani, 2019: Downscaling of rainfall extremes from satellite observations. *Water Resour. Res.*, **55**, 156–174, <https://doi.org/10.1029/2018WR022950>.
- Zubieta, R., M. Saavedra, J. C. Espinoza, J. Ronchail, J. Sulca, G. Drapeau, and J. Martin-Vide, 2019: Assessing precipitation concentration in the Amazon basin from different satellite-based data sets. *Int. J. Climatol.*, **39**, 3171–3187, <https://doi.org/10.1002/joc.6009>.