

Article

Improving the Regional Applicability of Satellite Precipitation Products by Ensemble Algorithm

Waseem Muhammad ^{1,2}, Hanbo Yang ¹, Huimin Lei ^{1,*} , Ajmal Muhammad ³ and Dawen Yang ¹

¹ State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China; engr.jatoi@hotmail.com (W.M.); yanghanbo@tsinghua.edu.cn (H.Y.); yangdw@tsinghua.edu.cn (D.Y.)

² Center of Excellence in Water Resource Engineering, University of Engineering and Technology, Lahore 54890, Pakistan

³ Department of Agricultural Engineering, University of Engineering and Technology, Peshawar 25120, Pakistan; engr_ajmal@uetpeshawar.edu.pk

* Correspondence: leihm@tsinghua.edu.cn; Tel.: +86-10-6278-3383

Received: 11 January 2018; Accepted: 3 April 2018; Published: 9 April 2018



Abstract: Satellite-based precipitation products (e.g., Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) and its predecessor, Tropical Rainfall Measuring Mission (TRMM)) are a critical source of precipitation estimation, particularly for a region with less, or no, hydrometric networking. However, the inconsistency in the performance of these products has been observed in different climatic and topographic diverse regions, timescales, and precipitation intensities and there is still room for improvement. Hence, using a projected ensemble algorithm, the regional precipitation estimate (RP) is introduced here. The RP concept is mainly based on the regional performance weights derived from the Mean Square Error (MSE) and the precipitation estimate from the TRMM product, that is, TRMM 3B42 (TR), real-time (late) (IT) and the research (post-real-time) (IR) products of IMERG. The overall results of the selected contingency table (e.g., Probability of detection (POD)) and statistical indices (e.g., Correlation Coefficient (CC)) signposted that the proposed RP product has shown an overall better potential to capture the gauge observations compared with the TR, IR, and IT in five different climatic regions of Pakistan from January 2015 to December 2016, at a diurnal time scale. The current study could be the first research providing preliminary feedback from Pakistan for global precipitation measurement researchers by highlighting the need for refinement in the IMERG.

Keywords: satellite observation; regional estimates; ensemble algorithm; TRMM

1. Introduction

The accurate quantification of precipitation is critical for significant flood projection, drought assessment and water resources management practices. Furthermore, in the context of current climate change, it is necessary to advance our understanding of temporal and spatial precipitation dynamics, both regionally and globally. The accuracy of the precipitation input also substantially influences the efficacy of hydrological models. Furthermore, obtaining an accurate regional precipitation database for significant hydrological predictions is essential; however, it is considered a challenging task in developing countries like Pakistan with complex terrain with no, or an insufficient, temporal database.

Traditionally, precipitation is measured through ground-based observation using several rain gauging networks/instruments, such as rain gauges, weather radar and/or interpolation of rainfall measured at rain gauges. However, many parts of the developing, and even the developed, world are

characterized by no, or sparse, density and inconvenient spatial distribution of hydrometric networking and therefore obtaining spatial rainfall observations are subject to even larger uncertainties.

To tame the aforementioned issue of ground-based rainfall observation, space-based precipitation products are used as a promising alternative source. Numerous, high-resolution satellite-based precipitation products are recently functioning to provide an unprecedented opportunity for a hydrologist to monitor the temporal and spatial distribution of precipitation on a near-global scale. The most widely used products include the Climate Prediction Center morphing technique (CMORPH) [1], the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)'s Multi-sensor Precipitation Estimate (MPE) [2], the European Centre for Medium-Range Weather Forecasts (ECMWF)'s Era-Interim product [3], the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), the PERSIANN Cloud Classification System estimation (PERSIANN-CCS) [4], the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis products (TMPA) versions 6 (3B42-V6) and 7 (3B42-V7) [5,6] and the Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) product [7]. Numerous studies have been made to evaluate the performance of these precipitation satellite products on the regional and global scale (e.g., Asia [8–11], North and Central America [12–16], South America [17–20], Europe [21–24], Australia [25–28], Oceans [29,30], other [31–36] and in Pakistan [37–41]).

Based on a comprehensive evaluation, these studies concluded that: (i) Version 7 of the TRMM product outperforms the preceding version 6 by substantial improvements in the bias; (ii) Based on the mean error, correlation with ground-based observation and the false alarm ratio, the TRMM 3B42 resulted in a significantly better performance compared to other satellite products across many regions [34]. This may be because of the bias adjustment integrated into the 3B42 algorithm. Though the 3B42 could be strongly biased in regions with sparse density and inconvenient spatial distribution of hydrometric networking; nonetheless, the product resulted in a low probability of detection in Australia and some parts of South America [31]; (iii) As compared to the TRMM, the IMERG showed an appreciably better performance across the mainland of China, however, no significant improvement was noted in the data-sparse mountainous watershed of Myanmar. Though satellite products, for example, the IMERG and its predecessor the TRMM version 3B42, could potentially capture the spatiotemporal variability of precipitation, however, it contains considerable error and there is still room for further improvement of the capturing capability for both products in a dry climate and high altitude constituency [33,34]; (iv) In complex terrain and in regions with rapid precipitation gradients, the rain detection efficacy of most of the satellite products showed a weak performance and mean errors mainly depended on magnitude. This may be because of the poor ability to distinguish between raining and non-raining clouds. Even though Global Satellite Mapping of Precipitation moving vector with Kalman filter (GSMaP-MVK) showed high correlations within situ data over Japan, there is still no universally acceptable product whose performance could be considered the best in these types of regions. Hence, over precipitous regions, further validation research could provide a better understanding of the limitations of using satellite products in flash flood applications [31]; (v) In addition to topographic diversity, the climatological features and seasonality also play a vital role in the better performance of these products, specifically in terms of mean errors and the probability of detection [32]. A significantly better performance was shown in the equatorial and tropical region as compared to the semi-arid region. Moreover, the summer (warm) season associated with the convective structure and cold seasons with light rainfall played a decisive part in the performance evaluation of these products [32]; (vi) In Pakistan, the TRMM products provide a better performance in plain and medium elevation areas, however, resulted in overestimation in the glaciated and mountainous areas in the north and in coastal areas and in arid regions, their use is questionable [38,41]. The error associated with the satellite products showed significantly geo-topographic dependability [37–41]; (vii) More conclusively, the performance of these products mainly based on the variability of many factors, for example, climatic and topographic

diversity, timescales (annual, seasonal, daily, and monthly), precipitation intensities, season (monsoon, winter etc.) [31,38,42] and, irrespective of significant improvements, in most of satellite precipitation products, still a region-specific assessment is necessary before any hydrological simulation, assessment, projection or outlook studies [42].

The current study introduces an ensemble algorithm based quantification of precipitation estimates considering a blended precipitation estimator. The basic idea behind the concept was to use the relative regional performance weights of the various well-known satellite precipitation products (here, IMERG research (IR), IMERG real-time (IT) and TRMM 3B42 (TR)) and to provide significant final regional precipitation estimates by exploiting the advantages and minimizing the disadvantages of these satellite products. More specifically, the main objectives of this study were; (i) To evaluate the performance of satellite products across the divergent climatic regions of Pakistan; (ii) To quantify blended precipitation estimates aiming at a better performance compared to the individual satellite observations in all climatic and topographic diverse regions of Pakistan. Additionally, providing the preliminary feedback from Pakistan for global precipitation measurement researchers by highlighting the need for refinement in the IMERG.

2. Materials and Methods

2.1. Study Area

Pakistan is a developing country and ranges from approximately 61–77°E (longitude) and 23.5–37°N (latitude) with a coverage of 79.6 Million ha area and an elevation ranging from 0 m (Arabian sea) to 8611 m (K2-Mount Godwin-Austin mountain). The total cultivated land is about 24% of the total country area, of which irrigated is about 80%, land covering Forests and grazing is about 4%, 31% is unfit for agriculture and 2% under cover. Landscape diversity ranging from coastline alongside the Arabian Sea in the south, plains area, deserts, and plateaus in the middle, to the snowed mountainous region in the north (Figure 1). The geographical and water resources based division includes; the Indus riverine covering a major area of Punjab, Khyber Pakhtunkhwa, and the Sind province of Pakistan, the Himalayan mountainous region (North–East), Northern Highland, and the drought-prone arid climatic region southwest region. Owing to these diverse climatic regimes, the significant spatiotemporal variation in precipitation (i.e., from 300 mm in the south to about 1500 mm in the north) occurs throughout the country. For the measurement of the precipitation, the ground-based hydrometric networking is installed, however, characterized by no or inadequate density and inconvenient temporal and spatial resolutions of rain gauge networking for hydrologic applications.

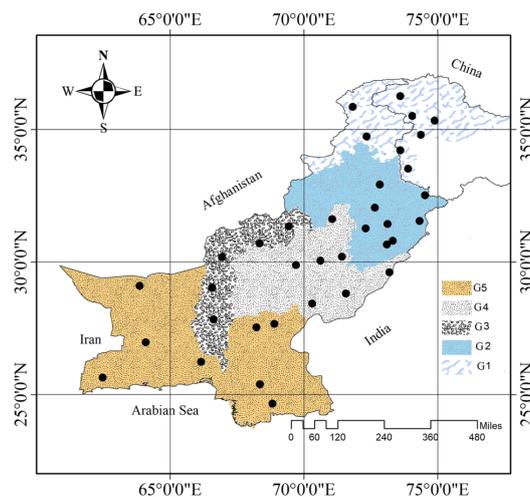


Figure 1. Spatial distribution of five different climatic region (G-1 to G-5) and Gauges in the selected area of study.

2.2. Datasets

2.2.1. Rain Gauges

The climate of Pakistan is in general characterized by a warm summer time and cold winters with a wide range of variations between extremes in different regions. To monitor this variation, currently, the Pakistan Meteorological Department (PMD) has established significant weather stations (including those from 1950, seasonal ones and new ones) all over Pakistan. Based on 30 years' precipitation data trend and climatic analysis of different regions, Chen et al. [43] had divided Pakistan into five different climatic regions (detailed below). Using this classification, we selected several stations from each climatic region (Figure 1) for the evaluation of selected satellite products and the quantification of proposed precipitation estimates. The preliminary selection of the thirty-five stations was to ensure the quality of reference data, independencies from gauge-adjusted satellite product, confirmed by the PMD, (source agency) and a continuous string of the observed data with no missing data. Furthermore, the salient features of the considered region are as follows; Region 1 (G-1): G1 includes a very cold climate in winter, a mild temperature in summer with high green mountains range, situated in the north (between 34°N to 38°N) of Pakistan. Region 2 (G-2): G2 has a mild cold climate in winter, a relatively hot summer, Sub Mountains, and located between 31°N to 34°N. Region (G-3): G3 comprises of relatively cold winters and hot in summer. Most of the stations have dry mountainous areas covering an area between 27°N to 32°N. Region (G-4): G4 is the hottest and driest region of the country where highest maximum temperatures (53 °C) are generally recorded in summer. The area is almost plain with some desert areas like Thal, Cholistan, and Thar. Region (G-5): G5 is a big region with a coastal range and is mostly arid to hyper-arid.

2.2.2. Satellite Based Estimates

Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG)

The IMERG V.04, level 3 is a quasi-global (60°N to 60°S) multi-satellite precipitation product of Global Precipitation Measurement (GPM) (<https://pmm.nasa.gov/GPM>), which provides estimates based on the combined use of passive microwave (PMW) sensors, infrared (IR) satellites, and ground-based precipitation data. The GPM-IMERG mission was launched in 2014 (February), aiming at continuing and improved satellite precipitation and snowfall products on a global scale. The level 3 products comprise of gridded precipitation and a snowfall database, with a spatial resolution of $0.1^\circ \times 0.1^\circ$ and temporal resolution of 30 min, estimated from the combined GPM constellation satellites, and calibrated by the Global Precipitation Climatology Centre (GPCC)'s gauge analysis [33,34,44].

The IMERG generally run twofold, first to generate the IMERG-Early/post real-time database (about 6 h after minimal observation time) for a warning from probable flood or landslides, and next to produce the IMERG-Late observation with about 18 h latency for studies of drought monitoring or agricultural forecasting. After receiving monthly gauge observations, the IMERG-Final cycle is run to generate the database approximate latency 3-months later, the observation month. In this study, we use the calibrated GIS version of daily real-time (IMERG-late) (IT) and research (IMERG-post-real-time) (IR) products to accomplish the projected objectives.

Tropical Rainfall Measuring Mission (TRMM)

The TRMM_3B42v7, Multi-satellite Precipitation Analysis (TMPA) (<https://pmm.nasa.gov/data-access/downloads/trmm>) launched in 1997, is envisioned to provide the best satellite estimate of precipitation on a nearly quasi-global scale between 50°S and 50°N with spatial resolution of $0.25^\circ \times 0.25^\circ$ and a temporal resolution of 3 h, accumulated daily and monthly using combined satellite precipitation related sensors. The TRMM combines IR radar with four PMW sensors named Microwave Imager (TMI), Precipitation Radar (PR), and Special Sensor Imager (SSI) and advances

Microwave Scanning Radiometer (AMSR). Here, the TRMM-TMPA 3B42 Real Time (hereafter TR) was used in order to evaluate its performance and using candidates to be an ensemble for the quantification of regional precipitation estimates in Pakistan.

2.3. Evaluation of Precipitation Products

As the selected satellite precipitation products differ in spatial resolution, initially an aggregating method was used to perform further analysis on the same spatial resolution (i.e., 0.25°) and daily timescale. Using the aggregation method, the daily precipitation products of the IMERG was aggregated from a 0.1° to 0.25° spatial resolution. For precipitation aggregation, we assigned areal weights, that is, 0.16, 0.08 and 0.04 to four IMERG grid cells falling completely inside a 0.25° TRMM grid cell, four located halfway within, and the ninth covers one-fourth inside, respectively [34]. For better data quality, the approximate similarity between the time of observations for both satellite (09.30 a.m. PST, +5 in case of UTC 1430) and gauge sources (08 a.m. PST, +5 in case of UTC 1300) insured. Furthermore, to avoid any mismatch issue, using the latitude/longitude of the rain gauges, the amount of satellite precipitation in the pixel against the exact location of each rain gauge was extracted.

Moreover, following the point to pixel analysis [45], the IMERG and TRMM V7 precipitation observations at the grid cell (box), where at least one weather station's falls were compared with the corresponding gauged observations from January 2015 to December 2016. To systematically assess the efficacy of IMERG and TRMM products, several statistical indices were used, which include: (a) The correlation coefficient (CC) (Equation (1)), to evaluate the agreement between the satellite products observations and gauge-based precipitation; and (b) The Mean Square Error (MSE), (Equation (2)) to assess the average absolute error.

In addition to the above two indices, the contingency of satellite products was represented using: (c) The probability of detection (POD)/hit rate, (Equation (3)) representing a fraction of correctly detected precipitation events by the satellite products among all the gauged-based precipitation events—moreover, 1 mm was used as the rain/no rain threshold [44]; (d) False alarm ratio (FAR), (Equation (4)) denotes the fraction of incorrectly detected events by satellite records rainfall; and (e) The Critical success index (CSI), (Equation (5)) shows the overall fraction of precipitation events correctly detected by the satellite products. It is also valuable to mention here that the ideal values of CC, POD, and CSI is one, while in the case of FAR the value is zero, whereas, the product with minimum MSE is considered comparatively reliable.

$$CC = \frac{\sum_{i=1}^N (S_i - \bar{S})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^N (S_i - \bar{S})^2} \sqrt{\sum_{i=1}^N (G_i - \bar{G})^2}} \quad (1)$$

$$MSE = \left[\frac{\sum_{i=1}^N (S_i - G_i)^2}{N} \right] \quad (2)$$

$$POD = \frac{hit}{hit + miss} \quad (3)$$

$$FAR = \frac{false..alarm}{hits + false..alarm} \quad (4)$$

$$CSI = \frac{hits}{hits + miss + false..alarm} \quad (5)$$

where, S_i and G_i represents daily satellite precipitation product in mm, and gauge-based observation for the i th time step, respectively \bar{S} and \bar{G} are average values (mm), N indicate the sample size of selected time series, *hits*: means the numbers of events when both satellite product and gauged based data record the precipitation, *miss*: indicating the events recorded by gauged based observation,

while missed by satellite product, false alarms: events when satellite product capture the precipitation, while gauged observation showing no precipitation.

2.4. Proposed Algorithm for Regional Satellite Precipitation Estimate Quantification

The proposed regional satellite precipitation estimate quantification framework is mainly based on the leave-one-out cross validation (LOOCV), regional performance weights of the considered satellite products, and the ensemble algorithm. In LOOCV, the data record of one station (assumed ungauged station) in a specific region was held out from the calibration database. Then the regional weights associated with each product were estimated using the calibration results of the remaining station (assumed donors). The LOOCV was repeated for all stations considered in the current study. The estimation process of regional weights and regional satellite precipitation estimate is as follows.

- i. To apply the RP algorithm in selected five (5) diverse regions, Let for a specific gauge in any of selected region, S^j (whereas $j = 1, 2, 3, \dots, n$) be the any available satellite product to be considered, $S_i^j \leftarrow [S_1, S_2, S_3, \dots, S_N]$ (mm) be the satellite precipitation observations time series of j th satellite product (here IR, IT, and TR) and $G_i \leftarrow [G_1, G_2, G_3, \dots, G_N]$ (mm) be the corresponding gauge-based observed time series of size N . Using G_i gauged observation and corresponding S_i^j time series of selected satellite products (TR, IT, IR), individual mean square error (Equation (6)) associated with each j th satellite product has been computed.

$$MSE^j = \frac{1}{N} \sum_{i=1}^N (G_i - S_i^j)^2 \quad (6)$$

- ii. Using estimated MSE^j of all donor gauged stations of a specific region (i.e., G_1 , or G_2 or, \dots , G_5), the regional MSE_r^j associated with each satellite product of that climatic region was estimated by simply averaging the MSE^j of all donor stations.
- iii. The product that does not provide the significant agreement with observed data can tend to produce high inaccuracy (high MSE). Hence, preliminary, instead of blindly combining the products, the S^j with comparatively highest MSE_r^j was eliminated through MSE filter from the further process. The intention behind the elimination was to minimize the future quantification overhead while enhancing accuracy. More specifically, the subset k (whereas $k = 1, 2, 3, \dots, m$) of satellite products with least MSE_r^j value has been retained by eliminating $j-k$ (here one) products having the highest MSE_r^j for further quantification.
- iv. Further, based on performance (MSE_r^j value), the weights were assigned to individual retained k th satellite product using Equation (7).

$$w_j = 1 / MSE_r^j \left(\sum_{j=1}^n \frac{1}{MSE_r^j} \right) \quad (7)$$

where $\sum w_j = 1$.

- v. After calculation of the respective regional weights of the best subset satellite products (k) (Figure 2), the satellite precipitation observations time series of k satellite products have been ensemble using Equation (8) to quantify the regional precipitation estimates (RP) for the assumed ungauged station.

$$RP_i = w_1 S_i^1 + w_2 S_i^2 + w_3 S_i^3 + \dots + w_k S_i^k \quad (8)$$

- vi. Lastly, the developed $RP_i \leftarrow [RP_1, RP_2, RP_3, \dots, RP_N]$ time series for a specific assumed ungauged station was evaluated by corresponding observations $G_i \leftarrow [G_1, G_2, G_3, \dots, G_N]$

using Equations (1)–(6) to validate the efficacy and projected improvement compared to individual selected j satellite products.

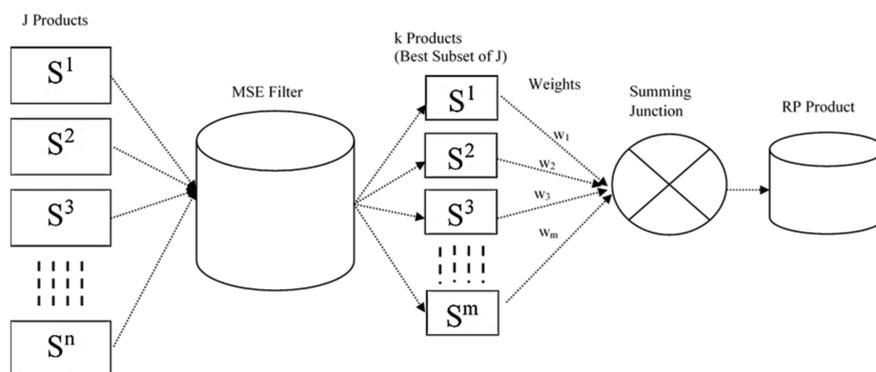


Figure 2. Flow diagram of the projected algorithm for the regional satellite precipitation estimate (RP).

3. Results

3.1. Spatial Variability of Product Performance at Grid-Scale

All diurnal satellite-based precipitation time series—i.e., TR, IR, IT data—against the rain-gauge data were compared at grid having at least one station. Preliminary the emphasis was on the evaluation of the performance of all satellite product at different climatic regions. The performance evaluation was carried out using various metrics i.e., CC, POD, FAR, CSI and MSE. The spatial variability of these metrics for TR, IR, and IT daily product over the selected five different climatic regions are shown in Figure 3. The CC based evaluation ensued that compared to IR and TR, IT agreed well with the majority of the gauge observations. However, the CC was little lower in most parts of the selected regions and even under 0.2 in few gauge stations. The POD was significantly comparable in the case of IT and TR; resulting in similar spatial characteristics in most of the cases; however, IR was slightly lower over most of the stations, whereas the FAR distribution ensued quite unlike result from the preceding metrics. Based on the FAR evaluation, the TR resulted in more significant outcomes with low FAR compared to IR and IT in all the regions and MSE shows a quite similar distribution in most of the cases. The performance comparison based on MSE noted that although these three satellite products produced substantial results at a specific gauge, none of the above performed consistently well at all gauges. Among the three selected products, the TR and IT performed better at the majority of the stations in G1 (with average MSE value of 81 and 97 mm²), G3 (with average MSE value of 10 and 10 mm²), and G4 (with avg. MSE value of 38 and 41 mm²). Whereas, the IR and TR showed significant results in G2 (with avg. MSE value of 108 and 76 mm²) and G5 (with average MSE value of 14 and 17 mm²) (see Figures 4 and 5). There was no particular reason to prefer, a priori, one of these for all stations. Hence, to assess consistency and reliability throughout the entire study region, the projected RP, based on the ensemble algorithm, was also analyzed here.

The same metrics for performance evaluation were adopted to validate the hypothesized framework. Table 1 shows the complete Evaluation Statistic for the PR algorithm at each selected gauge. The results (Table 1 and Figure 3 (RP-section)) illustrated that the CC and POD significantly increased as compared to the TR, IR, and IT throughout the study area. A very few stations (having less than five rainfall events) still evidenced the low values of CC and POD. Furthermore, to test whether the difference in the evaluation measures (CC, POD etc.) are statistically significant or not, we performed a nonparametric test (i.e., Wilcoxon signed rank sum test) using the sample size $n = 36$. The Wilcoxon signed rank sum test was performed to test whether our research finding, that is, that the RP performed better compared to other satellite products (alternative hypothesis H1), is meaningful or whether there is no significant difference in outcomes (Null hypothesis Ho), at a significance level of $\alpha = 0.05$.

The Wilcoxon signed rank sum test showed that there is a significant difference between the RP and all selected satellite products (i.e., TR, IR and IT) in terms of POD with a p value less than α , a Wilcoxon test statistic (w) less than the critical value (i.e., $w_{critical}(for.n = 36) = 227$ and an absolute z-statistic $|z| > z_{critical}$. Moreover, in the case of CC, a significant statistical difference was also observed between the RP and IR [$|z|(4.861) > z_{critical}(1.96)$ and $w(0) \leq w_{critical}(227)$], RP and TR [$w(39.5) \leq w_{critical}(227)$], and between RP and IT [$w(64.5) \leq w_{critical}(227)$].

Figure 4 shows the competitive performance evaluation of the IR, IT, TR, and RP over all the selected stations in Pakistan. The CC, POD, MSE and CSI based evaluation result was that the RP product significantly outperformed a majority of other stations (as shown with the blue circles) and provided a consistent performance in all the diverse regions of Pakistan. It may be worthy here to mention that low performance value of RP can be resulted in case of poor performance of all the TR, IR and IT at the specific station.

Table 1. Evaluation Statistic for the precipitation radar (PR) algorithm at each selected gauge.

Evaluation Statistic for the PR Algorithm							
Station #	Lat	Long	CC	POD	CSI	FAR	MSE
1	24.6	68.9	0.68	0.70	0.19	0.21	21.03
2	29.9	73.3	0.49	0.59	0.26	0.32	30.86
3	29.3	71.8	0.41	0.74	0.32	0.36	24.53
4	29.9	69.7	0.22	0.67	0.33	0.39	36.09
5	31.6	71.1	0.46	0.66	0.29	0.35	44.60
6	25.1	61.8	0.51	0.71	0.24	0.26	4.80
7	32.9	72.8	0.18	0.68	0.29	0.34	72.81
8	35.4	74.1	0.31	0.37	0.16	0.22	13.58
9	35.9	71.8	0.20	0.41	0.24	0.37	21.46
10	30.1	70.6	0.41	0.68	0.32	0.37	26.26
11	31.4	73.1	0.24	0.55	0.26	0.32	51.72
12	35.9	74.3	0.26	0.28	0.17	0.30	9.44
13	25.1	62.3	0.53	0.70	0.32	0.37	2.96
14	25.4	68.4	0.21	0.77	0.19	0.20	22.08
15	31.3	72.3	0.35	0.54	0.25	0.32	57.72
16	29.0	66.6	0.49	0.45	0.09	0.10	1.84
17	27.8	66.6	0.14	0.41	0.19	0.27	8.66
18	33.5	73.9	0.44	0.65	0.41	0.52	115.85
19	35.3	74.9	0.29	0.52	0.32	0.45	22.07
20	31.6	74.3	0.23	0.59	0.34	0.44	101.32
21	27.5	68.2	0.16	0.52	0.21	0.27	5.34
22	26.2	66.2	0.37	0.77	0.22	0.23	8.35
23	30.2	71.4	0.53	0.66	0.30	0.35	12.38
24	34.4	73.5	0.38	0.67	0.45	0.57	91.16
25	34.2	73.6	0.32	0.68	0.44	0.55	99.88
26	30.8	73.4	0.28	0.61	0.27	0.33	55.90
27	27.0	64.1	0.34	0.68	0.22	0.25	2.31
28	30.7	68.4	0.14	0.48	0.24	0.32	14.78
29	30.2	67.0	0.24	0.32	0.15	0.22	6.74
30	28.4	70.3	0.19	0.57	0.17	0.20	17.65
31	30.6	73.2	0.13	0.64	0.29	0.35	52.95
32	32.0	72.7	0.33	0.56	0.30	0.39	87.47
33	32.5	74.5	0.53	0.72	0.39	0.46	112.38
34	27.7	68.9	0.24	0.48	0.18	0.23	9.10
35	34.7	72.3	0.29	0.55	0.32	0.43	66.39
36	30.1	67.0	0.24	0.38	0.17	0.24	9.54

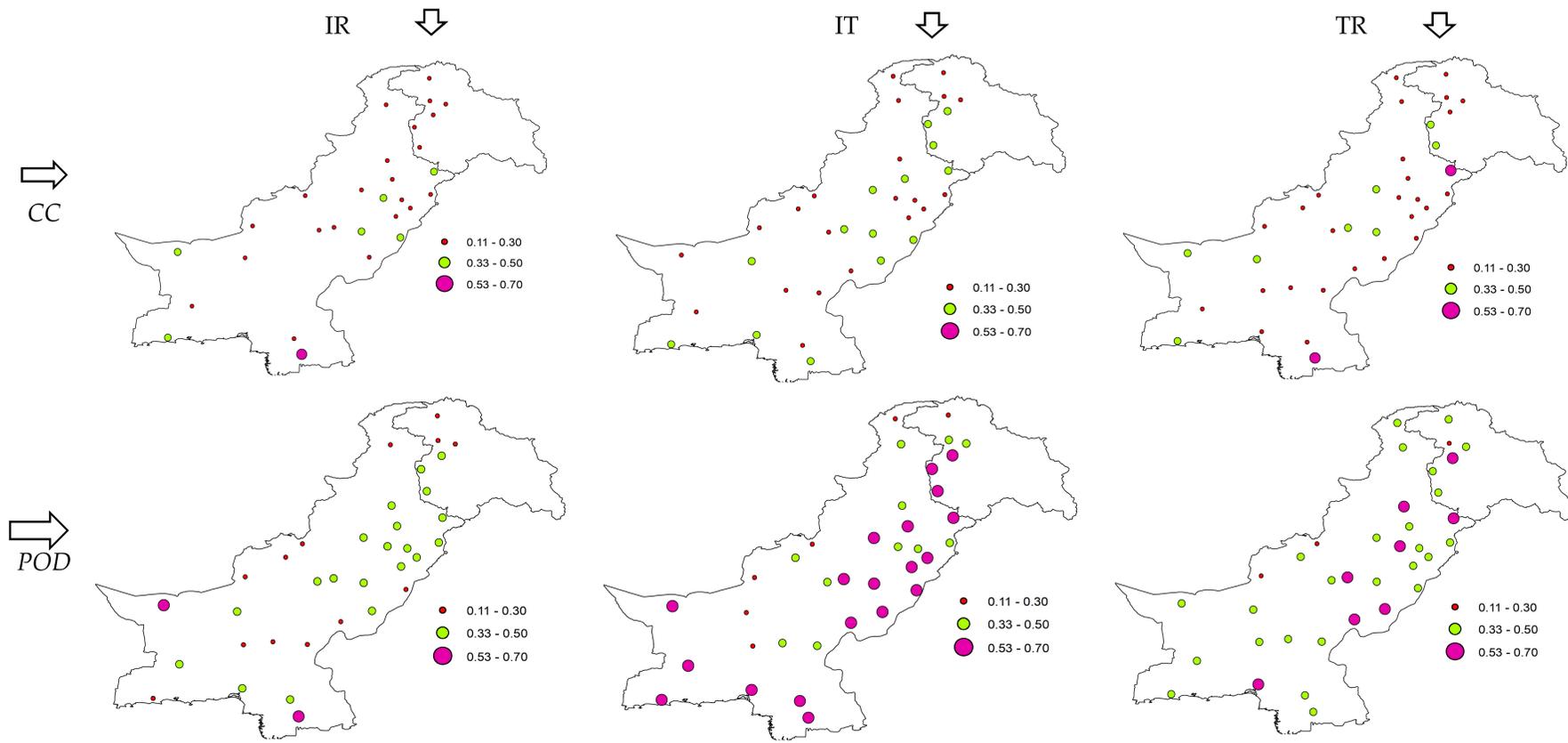


Figure 3. Cont.

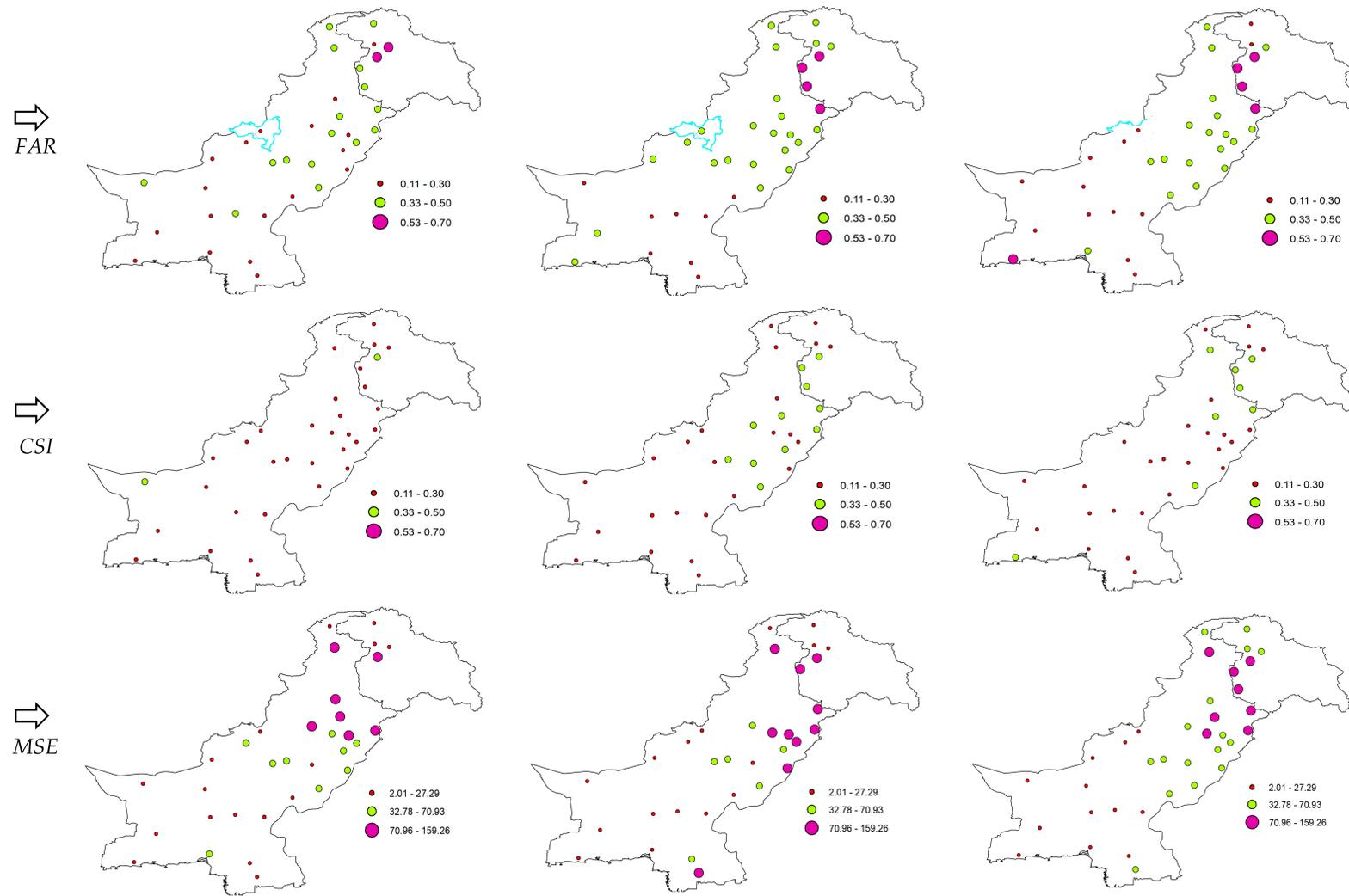


Figure 3. Cont.

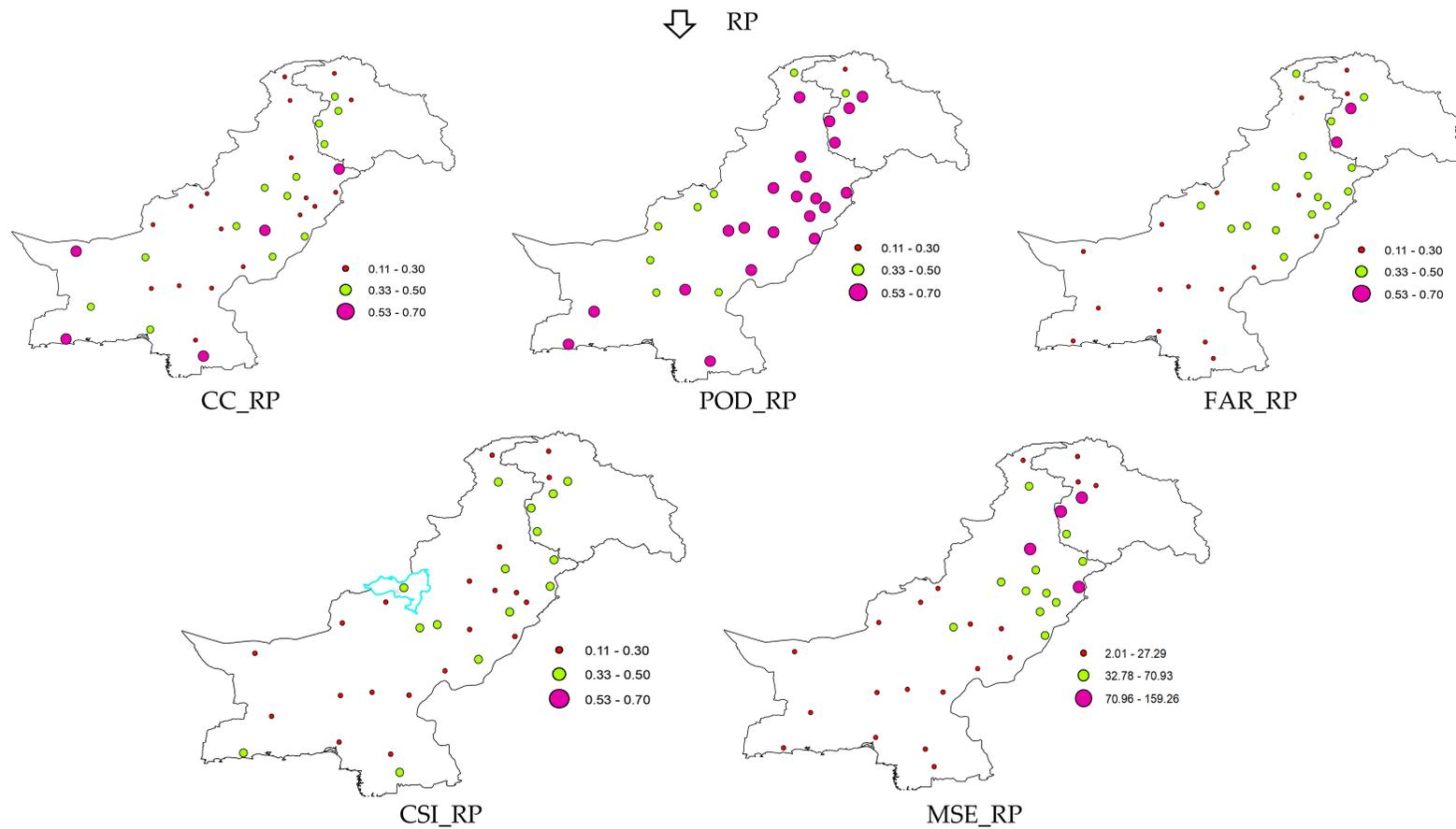


Figure 3. Spatial distribution of the selected metrics (i.e., correlation coefficient (CC), probability of detection (POD), False alarm ratio (FAR) and Critical success index (CSI) and mean square error (MSE) (in mm^2) for the IMERG research (IR), IMERG real-time (IT) and TRMM 3B42 (TR) and regional precipitation (RP) over Pakistan (Note: above RP: Column shows Satellite Product, while row indicating metrics, below RP: distribution of all metrics for the RP only).

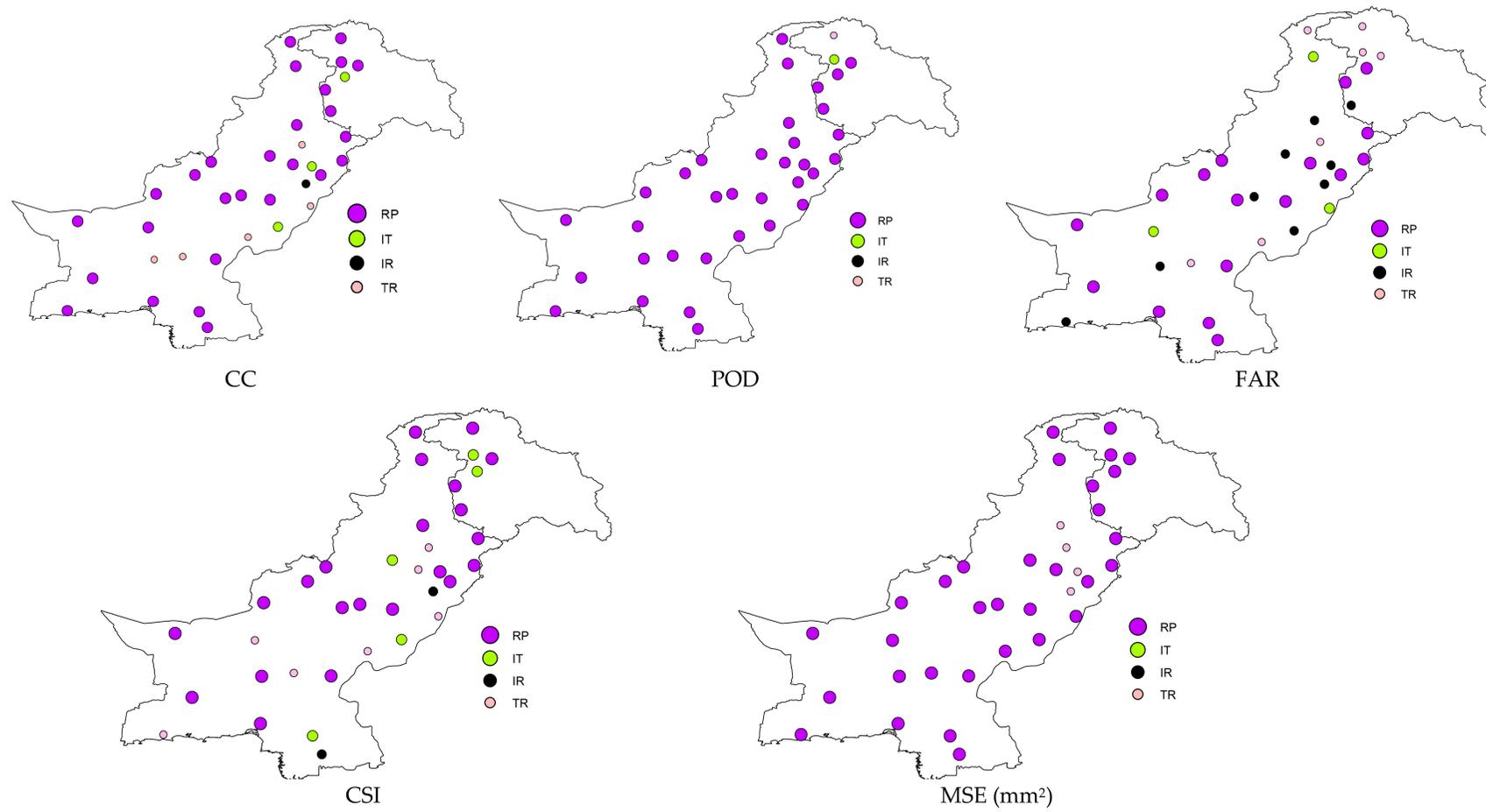


Figure 4. Comparative performance assessment for the RP, IR, IR, and TR, at each station.

The box plot (Figure 5) provides the graphical summary for the performance of each satellite product. The accuracy of the products can be observed by the minimum (lower line) first quartile 3rd quartile (central rectangle), whisker (segment inside the rectangle) and maximum (top small line) of the given box plots drawn against each product for whole study areas. The statistical summary indicates that the RP provides a comparatively better performance based on all the metrics.

3.2. Spatial Distribution of Product Performance at Regional Scale

To further explore the continuity, consistency of the TR, IR, IT and RP precipitation estimate, the products were evaluated and compared at the regional scales (Figure 6) using the evaluation metrics over the five different climatic regions. The metrics evaluation ensured that in most cases, the RP outperformed the other products in term of CC, CSI, FAR and more specifically in terms of POD and MSE. The POD and MSE based performance evaluations show that the performance has been improved greatly compared to either product. The higher POD and lower MSE value in all regions indicated that, irrespective of the region, the RP could perform better compared to other selected satellite products. It was expected, as the RP significantly exploits the advantages and minimizes the disadvantages of these satellite products. An outcome-based conclusion resulting in a significant reliability of agreement has been observed between the RP and gauged observations. Hence, such a blended approach could be applied to pixels without gauges and considered a step forward in approving the efficacy of satellite-based observations at the regional scale with no, or less, hydrometric networking.

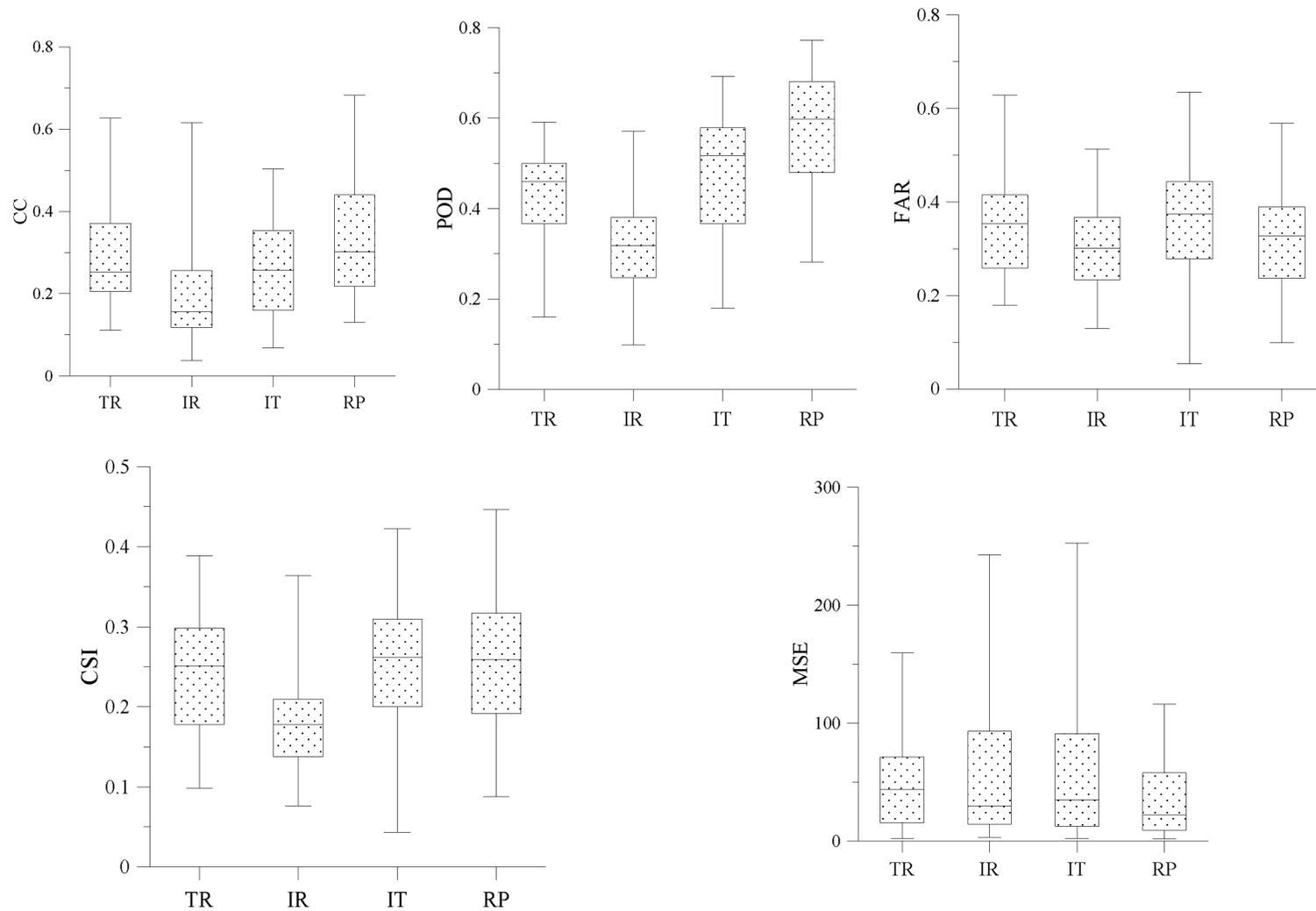


Figure 5. Comparative statistical summary of CC, POD, FAR, CSI, and MSE (mm²) based performance of the TR, IR, IT, and RP.

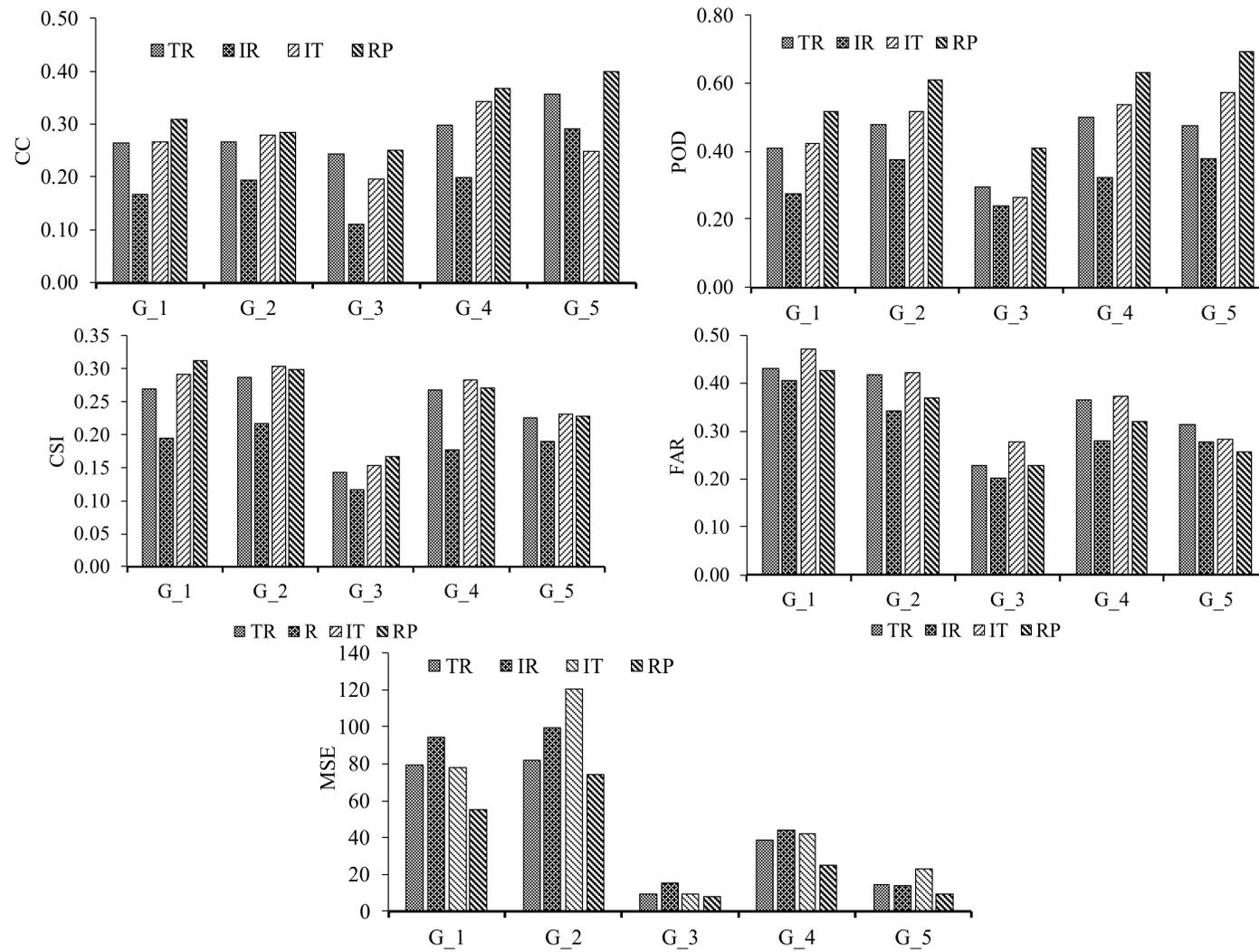


Figure 6. Comparative performance assessment based on the average value of considered evaluation metrics.

4. Discussion

Satellite-based precipitation products are promising alternative sources of observation in a region with sparse hydrometric networking. Several studies have been carried out to develop and evaluate these products in Pakistan [37–41] and all over the world. However, the supremacy of any specific product has not been observed in different regions with diverse climatic conditions [31,42]. In the current study, the quantitative comparison between three satellite products' estimates and projected precipitation estimates was carried out using gauge-based observations. The analysis outcomes presented in Figures 3–6 support the finding of evolutionary studies [37–41] in Pakistan. The results demonstrated that the selected satellite product (TR, IR, and IT) had significant potential to capture the gauge observation; however, performance was inconsistent in different regions of Pakistan. Similar to the findings of [38,41], the TMPA product provided substantial outcomes in term of CC, FAR, POD, CSI and MSI- in low altitude regions, however, in the case of high altitude, glaciated regions, dry areas and regions with fewer events of rainfall, the performance declined. Anjum et al. [39] evaluated real-time and post real-time TRMM precipitation estimates for a single heavy precipitation event (28–30 July) over a watershed and results in low correlation and underestimation. Furthermore, Cheema et al. [40] also attempted to calibrate the TRMM product at a basin scale (Indus basin) and different temporal scales by using regression and geographical differential analysis. More specifically, Anjum et al. and Cheema et al. [39,40] also conceded the geo-topographic, climatic dependability, and low performance of these products in Pakistan, which also second the outcomes of the current study. There could be numbers of reason, for example, uncertainty in verification of results in developing countries like Pakistan (e.g., station density, the impact of wind flow, random static error), complex topography and climate over a specific region, seasonality (monsoon, pre-monsoon, post-monsoon, winter, summer etc.) [31,38,42], and scarce use of the gauge in the production of GPM gauge observation estimation etc.

To avoid this kind of non-negligible error in these satellite products, several efforts have been made, for example, improvement in calibration algorithms, moving from TRMM to IMERG, reducing sampling issue and so forth, however, still there is room for improvement to provide consistent results in diverse regions [33,34]. This study demonstrated that by merging different satellite products, there could be the possibility to minimize the inconstancy issue and the error associated with these products. The comparative outcomes second the hypothesis and the results showed that, despite the low performance (i.e., $CC_{avg} = 0.33$), the project RP framework was able to provide better agreement evaluation (i.e., low MSE_{avg} : 37.79 and high POD_{avg} : 0.61) with the gauged observation throughout the study area. This may be because of the ensemble algorithm used in the projected framework, which has the advantage of selecting a suitable subset of products based on performance and considering the final product. In addition, the plausible reason behind the low performance is owed to the total dependability on the performance of the candidate products.

Moreover, to the authors' knowledge, the IMERG is yet to be evaluated in the selected study area. Hence, besides providing a projected framework, the current study could be the first providing useful feedback from Pakistan for GPM researchers by highlighting the need for refinement in the IMERG products in Pakistan. In addition, it accentuates the need for denser and more reliable long-term precipitation observations networking to facilitate effective evaluation in the selected area of study.

5. Conclusions

In this study, the quality of different satellite products, that is, TRMM 3B42 (TR) and real-time (late) (IT) and research (post-real-time) (IR) products of IMERG have been evaluated over five different climatic and topographic regions of Pakistan. In addition to evaluation, an ensemble algorithm based on regional precipitation estimates (RP) have been introduced aiming at a better performance throughout the regions. The considered TR, IR, IT, and RP products were compared based on five distinct performance evaluation metrics (i.e., correlation coefficient (CC), Mean Square Error (MSE),

probability of detection (POD), False alarm ratio (FAR) and Critical success index (CSI) on the diurnal time scale. The outcomes-based main conclusions drawn are as follows:

- (i) Despite the inconsistency in different climatic regions, the TR, IR, and IT have shown significant potential to capture the gauge observations. Among these products, the IT outperformed the others at the majority of the stations; however, in G3 and G5, a slightly low performance was observed
- (ii) Although these products have performed well, still they contain considerable errors and there is a room for further improvement in their ability to provide better agreement throughout different regions
- (iii) The comparison ensured that the proposed RP framework provided significant agreement with the gauged observations throughout the study area and better outcomes. In addition, the proposed method has an advantage of selecting more suitable products based on performance and considering it as the final observation, therefore it can be considered a step forward in the improvement of the efficacy of the satellite-based observations
- (iv) As the projected framework is totally dependent on the performance of the candidate products, hence this total reliability could be a factor to be considered as its major drawback.

Acknowledgments: The National Key Research and Development Program of China (2016YFC0402404) and the National Natural Science Foundation of China (Project No. 51679120) financially supported the study. The authors would also like to thank the developers of the GPM and TRMM products.

Author Contributions: Waseem Muhammad, Huimin Lei, and Dawen Yang conceived and designed the experiments; Waseem Muhammad and Muhammad Ajmal performed the experiments; Waseem Muhammad and Muhammad Ajmal analyzed the data; Waseem Muhammad, Hanbo Yang, and Huimin Lei interpreted the results; Waseem Muhammad wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Joyce, R.J.; Xie, P. Kalman filter-based CMORPH. *J. Hydrometeorol.* **2011**, *12*, 1547–1563. [[CrossRef](#)]
2. Heinemann, T.; Kerényi, J. The EUMETSAT Multi Sensor Precipitation Estimate (MPE): Concept and Validation. In Proceedings of the 2nd International Precipitation Working Group (IPWG) Workshop, Monterey, CA, USA, 25–28 October 2004.
3. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [[CrossRef](#)]
4. Hsu, K.; Gao, X.; Sorooshian, S.; Gupta, H.V. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor.* **1997**, *36*, 1176–1190. [[CrossRef](#)]
5. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.; Hong, Y.; Bowman, K.P.; Stocker, E.F. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [[CrossRef](#)]
6. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Nelkin, E.J. *The TRMM Multi-Satellite Precipitation Analysis (TMPA): Satellite Rainfall Applications for Surface Hydrology*; Gebremichael, M., Hossain, F., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp. 3–22.
7. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Day 1 IMERG Final Run Release Note*; NASA/GSFC: Greenbelt, MD, USA, 2015.
8. Brown, J.E.M. An analysis of the performance of hybrid infrared and microwave satellite precipitation algorithms over India and adjacent regions. *Remote Sens. Environ.* **2006**, *101*, 63–81. [[CrossRef](#)]
9. Derin, Y.; Yilmaz, K.K. Evaluation of multiple satellite based precipitation products over complex topography. *J. Hydrometeorol.* **2014**, *15*, 1498–1516. [[CrossRef](#)]
10. Gao, Y.C.; Liu, M.F. Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 837–849. [[CrossRef](#)]

11. Guo, H.; Chen, S.; Bao, A.; Hu, J.; Gebregiorgis, A.S.; Xue, X.; Zhang, X. Inter-comparison of high-resolution satellite precipitation products over central Asia. *J. Remote Sens.* **2015**, *7*, 7181–7211. [[CrossRef](#)]
12. Anagnostou, E.N.; Maggioni, V.; Nikolopoulos, E.I.; Meskele, T.; Hossain, F.; Papadopoulos, A. Benchmarking high-resolution global satellite rainfall products to radar and rain gauge rainfall estimates. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1667–1683. [[CrossRef](#)]
13. Gourley, J.J.; Hong, Y.; Flamig, Z.L.; Wang, J.; Vergara, H.; Anagnostou, E.N. Hydrologic evaluation of rainfall estimates from radar, satellite, gauge, and combinations on Ft. Cobb basin, Oklahoma. *J. Hydrometeor.* **2011**, *12*, 973–988. [[CrossRef](#)]
14. Habib, E.; ElSaadani, M.; Haile, A.T. Climatology focused evaluation of CMORPH and TMPA satellite rainfall products over the Nile basin. *J. Appl. Meteor. Climatol.* **2012**, *51*, 2105–2121. [[CrossRef](#)]
15. Maggioni, V.; Vergara, H.J.; Anagnostou, E.N.; Gourley, J.J.; Hong, Y.; Stampoulis, D. Investigating the applicability of error correction ensembles of satellite rainfall products in river flow simulations. *J. Hydrometeor.* **2013**, *14*, 1194–1211. [[CrossRef](#)]
16. Qiao, L.; Hong, Y.; Chen, S.; Zou, C.B.; Gourley, J.J.; Yong, B. Performance assessment of the successive version 6 and version 7 TMPA products over the climate-transitional zone in the southern Great Plains, USA. *J. Hydrol.* **2014**, *513*, 446–456. [[CrossRef](#)]
17. Zulkafli, Z.; Buytaert, W.; Onof, C.; Manz, B.; Tarnavsky, E.; Lavado, W.; Guyot, J.L. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean–Amazon River basins. *J. Hydrometeor.* **2014**, *15*, 581–592. [[CrossRef](#)]
18. Salio, P.; Hobouchiand, M.P.; Skabarc, Y.G.; Vilaf, D. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmos. Res.* **2015**, *163*, 146–161. [[CrossRef](#)]
19. Falck, A.; Maggioni, V.; Tomasella, J.; Vila, D.A.; Diniz, F.L.R. Propagation of satellite precipitation uncertainties through a distributed hydrologic model: A case study in the Tocantins–Araguaia basin in Brazil. *J. Hydrol.* **2015**, *527*, 943–957. [[CrossRef](#)]
20. Mantas, V.M.; Liu, Z.; Caro, C.; Pereira, A.J.S. Validation of TRMM Multi-satellite Precipitation Analysis (TMPA) products in the Peruvian Andes. *Atmos. Res.* **2015**, *163*, 132–145. [[CrossRef](#)]
21. Stampoulis, D.; Anagnostou, E.N.; Nikolopoulos, E.I. Assessment of high-resolution satellite-based rainfall estimates over the Mediterranean during heavy precipitation events. *J. Hydrometeor.* **2013**, *14*, 1500–1514. [[CrossRef](#)]
22. Lo Conti, F.; Hsu, K.L.; Noto, L.V.; Sorooshian, S. Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea. *Atmos. Res.* **2014**, *138*, 189–204. [[CrossRef](#)]
23. Mei, Y.; Anagnostou, E.N.; Nikolopoulos, E.I.; Borga, M. Error analysis of satellite precipitation products in mountainous basins. *J. Hydrometeor.* **2014**, *15*, 1778–1793. [[CrossRef](#)]
24. Skok, G.; Zagar, N.; Honzak, L.; Zabkar, R.; Rakovec, J.; Ceglar, A. Precipitation intercomparison of a set of satellite- and raingauge-derived datasets, ERA Interim reanalysis, and a single WRF regional climate simulation over Europe and the North Atlantic. *Theor. Appl. Climatol.* **2016**, *123*, 217–232. [[CrossRef](#)]
25. Chen, Y.; Ebert, E.E.; Walsh, K.J.E.; Davidson, N.E. Evaluation of TMPA 3B42 daily precipitation estimates of tropical cyclone rainfall over Australia. *J. Geophys. Res. Atmos.* **2013**, *118*, 11966–11978. [[CrossRef](#)]
26. Peña-Arancibia, J.L.; van Dijk, A.I.J.; Renault, L.J.; Mulligan, M. Evaluation of precipitation estimation accuracy in reanalyses, satellite products, and an ensemble method for regions in Australia and South and East Asia. *J. Hydrometeor.* **2013**, *14*, 1323–1333. [[CrossRef](#)]
27. Pipunic, R.; Ryu, D.; Costelloe, J.; Su, C.H. Evaluation of real-time satellite rainfall products in semi-arid/arid Australia. In Proceedings of the 20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December 2013; pp. 3106–3112.
28. Pipunic, R.; Ryu, D.; Costelloe, J.; Su, C.H. An evaluation and regional error modeling methodology for near-real-time satellite rainfall data over Australia. *J. Geophys. Res. Atmos.* **2015**, *120*, 10767–10783. [[CrossRef](#)]
29. Prakash, S.; Gairola, R.M. Validation of TRMM-3B42 precipitation product over the tropical Indian Ocean using rain gauge data from the RAMA buoy array. *Theor. Appl. Climatol.* **2014**, *115*, 451–460. [[CrossRef](#)]
30. Klepp, C. The oceanic shipboard precipitation measurement network for surface validation—Ocean rain. *Atmos. Res.* **2015**, *163*, 74–90. [[CrossRef](#)]

31. Maggioni, V.; Meyers, P.C.; Robinson, M.D. A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. *J. Hydrometeorol.* **2016**, *17*, 1101–1117. [[CrossRef](#)]
32. Ringard, J.; Becker, M.; Seyler, F.; Linguet, L. Temporal and Spatial Assessment of Four Satellite Rainfall Estimates over French Guiana and North Brazil. *Remote Sens.* **2015**, *7*, 16441–16459. [[CrossRef](#)]
33. Tang, G.; Ma, Y.; Long, D.; Zhong, L.; Hong, Y. Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *J. Hydrol.* **2016**, *533*, 152–167. [[CrossRef](#)]
34. Yuan, F.; Zhang, L.; Win, K.W.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Assessment of GPM and TRMM Multi-Satellite Precipitation Products in Stream flow Simulations in a Data-Sparse Mountainous Watershed in Myanmar. *Remote Sens.* **2017**, *9*, 302. [[CrossRef](#)]
35. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.J.; Weedon, G.P.; Brocca, L.; Pappenberger, F.; Huffman, G.J.; Wood, E.F. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6201–6217. [[CrossRef](#)]
36. Gebregiorgis, A.S.; Hossain, F. How well can we estimate error variance of satellite precipitation data around the world? *Atmos. Res.* **2015**, *154*, 39–59. [[CrossRef](#)]
37. Khan, S.I.; Hong, Y.; Gourley, J.J.; Khattak, M.U.K.; Yong, B.; Vergara, H.J. Evaluation of three high-resolution satellite precipitation estimates: Potential for monsoon monitoring over Pakistan. *Adv. Space Res.* **2014**, *54*, 670–684. [[CrossRef](#)]
38. Iqbal, M.F.; Athar, H. Validation of satellite-based precipitation over diverse topography of Pakistan. *Atmos. Res.* **2018**, *201*, 247–260. [[CrossRef](#)]
39. Anjum, M.N.; Ding, Y.; Shangguan, D.; Ijaz, M.W.; Zhang, S. Evaluation of high-resolution satellite-based real-time and post-real-time precipitation estimates during 2010 extreme flood event in Swat River Basin, Hindukush region. *Adv. Meteorol.* **2016**, *2016*, 2604980. [[CrossRef](#)]
40. Cheema, M.J.M.; Bastiaanssen, W.G. Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *Int. J. Remote Sens.* **2012**, *33*, 2603–2627. [[CrossRef](#)]
41. Hussain, Y.; Satgé, F.; Hussain, M.B.; Martinez-Carvajal, H.; Bonnet, M.-P.; Cárdenas-Soto, M.; Roig, H.L.; Akhter, G. Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theor. Appl. Climatol.* **2018**, *131*, 1119–1132. [[CrossRef](#)]
42. Zambrano-Bigiarini, M.; Nauditt, A.; Birkel, C.; Verbist, K.; Ribbe, L. Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrol. Earth Syst. Sci.* **2016**, *21*, 1295–1320. [[CrossRef](#)]
43. Salma, S.; Rehman, S.; Shah, M.A. Rainfall Trends in Different Climate Zones of Pakistan. *Pak. J. Met.* **2012**, *9*, 37–47.
44. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China. *Remote Sens.* **2016**, *8*, 472. [[CrossRef](#)]
45. Thiemig, V.; Rodrigo, R.; Mauricio, Z.-B.; Vincenzo, L.; Roo, A.D. Validation of satellite-based precipitation products over sparsely gauged African river basins. *J. Hydromet.* **2012**, *13*, 1760–1783. [[CrossRef](#)]

