

Article

Impacts of Reference Precipitation on the Assessment of Global Precipitation Measurement Precipitation Products

Ye Zhang ¹, Leizhi Wang ^{1,*}, Yilan Li ^{2,3}, Yintang Wang ¹, Fei Yao ⁴ and Yiceng Chen ³¹ National Key Laboratory of Water Disaster Prevention, Nanjing Hydraulic Research Institute, Nanjing 210029, China² Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16802, USA³ Wuhan Zhongwei Intelligent Survey Technology Co., Ltd., Wuhan 430030, China⁴ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430070, China

* Correspondence: wanglz@nhri.cn

Abstract: Reference precipitation (RP) serves as a benchmark for evaluating the accuracy of precipitation products; thus, the selection of RP considerably affects the evaluation. In order to quantify this impact and provide guidance for RP selection, three interpolation methods, namely inverse distance weighting (IDW), ordinary kriging (OK), and geographical weighted regression (GWR), along with six groups of station densities, were adopted to generate different RPs, based on the super-high-density rainfall observations as true values, and we analyzed the errors of different RPs and the impacts of RP selection on the assessment of GPM precipitation products. Results indicate that the RPs from IDW and GWR both approached the true values as the station density increased ($CC > 0.90$); while the RP from OK showed some differences ($CC < 0.80$), it was similar to GWR when the station density was low, but the accuracy improved at first and then worsened as the station density continued to increase; the evaluation results based on different RPs showed remarkable differences even under the same conditions; when the average distance between rainfall gauges that were utilized to generate RPs was below the medium value (i.e., $d < 20$ km), the evaluation based on RP derived from IDW and GWR was close enough to that based on the true precipitation, which indicates its feasibility in evaluating satellite precipitation products.

Keywords: GPM; near-real-time satellite precipitation; reference precipitation; error analysis; accuracy evaluation



Academic Editors: Dong Ren, Huiqin Ma, Hang Sun and Li Liu

Received: 29 December 2024

Revised: 27 January 2025

Accepted: 9 February 2025

Published: 12 February 2025

Citation: Zhang, Y.; Wang, L.; Li, Y.; Wang, Y.; Yao, F.; Chen, Y. Impacts of Reference Precipitation on the Assessment of Global Precipitation Measurement Precipitation Products. *Remote Sens.* **2025**, *17*, 624. <https://doi.org/10.3390/rs17040624>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid advancement of space-based observation technologies, the ability of satellite-borne sensors to capture large-scale and even global-scale precipitation distributions has significantly improved. Since the launch of the Tropical Rainfall Measuring Mission (TRMM) in 1998, a series of global satellite-based precipitation products, including the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN [1]), the Climate Prediction Center MORPHing technique (CMORPH [2]), and the TRMM Multi-satellite Precipitation Analysis (TMPA [3]). In 2014, the Global Precipitation Measurement (GPM) mission, equipped with a multi-frequency microwave radiometer and dual-frequency radar, began providing more accurate references for calibrating multi-satellite precipitation retrievals [4]. This development marked the shift from the TRMM era to the GPM era in satellite precipitation product development.

Evaluating and analyzing the accuracy of precipitation products from GPM and other satellites has become a fundamental task in global quantitative precipitation estimation. Extensive research has been conducted globally on this subject [5], with a particular focus on the improvements in accuracy and applicability of satellite precipitation products from the TRMM to the GPM era [6,7].

Studies have indicated that GPM data, due to the inclusion of additional data sources and improvements in precipitation retrieval algorithms, perform exceptionally well in detecting heavy rainfall. However, existing research results often conflict with each other. Libertino et al. [8] evaluated the ability of the Integrated Multi-satellite Retrievals for GPM (IMERG) to detect heavy precipitation events on a global scale and found significant improvements compared to TMPA 3B42RTV7 data. Tang et al. [5] compared the monitoring and early warning capabilities of IMERG, the Global Satellite Mapping of Precipitation (GSMaP), TRMM, and PERSIANN products for the torrential rain-induced floods in southern China during the summer of 2016. Their results indicated that the IMERG and GSMaP products demonstrated significantly higher overall accuracy than the TRMM products. However, several studies have highlighted that GPM's ability to detect heavy rainfall is still unsatisfactory, and in some cases, even inferior to data from the TRMM era. For instance, Sahlu et al. [9] found that while the accuracy of the IMERG daily precipitation data is slightly better than that of CMORPH, its performance in detecting heavy rainfall is notably weaker. Additionally, studies by He et al. [10], Hayashi et al. [11], and Mekonnen et al. [12] also identified limitations in GPM's heavy rainfall detection capabilities. The conflicting conclusions from these studies may be attributed to the subjectivity in selecting reference precipitation (RP) during evaluation, which can significantly influence the results.

Evaluating and validating the accuracy of satellite precipitation products from GPM and other satellites requires RP to serve as the true value. Currently, there are two main types of RP. The first approach involves directly using precipitation observations from rainfall gauges within the satellite's precipitation grid cells as ground-based reference data [8,13,14]. If multiple gauges are present in a grid cell, the average or maximum rainfall value from these gauges is typically used as the reference. However, station observations represent point-scale precipitation, whereas satellite data reflect area-wide rainfall (e.g., GPM products have a spatial resolution of $0.1^\circ \times 0.1^\circ$). This creates a spatial scale mismatch between station measurements and satellite grid cells, which inevitably affects the reliability of evaluation results [15]. To address this issue, many studies have adopted a second approach, which involves using spatial interpolation methods to transform station observations into grid cells that align with satellite precipitation data [16–18]. While this method ensures spatial consistency between RP and satellite data, it also introduces interpolation errors [19–22]. Moreover, the selection of interpolation methods is often highly subjective, leading to variations in the RP generated by different techniques. Additionally, the quality of the precipitation dataset (e.g., the number of ground-based rainfall gauges) is another crucial factor influencing RP accuracy. Currently, there are very few specific studies examining the impact of RP generated by different methods on the evaluation of near-real-time satellite precipitation products. Therefore, quantitative research in this area is essential.

In light of these challenges, this study conducts a quantitative analysis of the impact of RP selection on the evaluation of satellite precipitation products. Four commonly used types of RP are generated based on the available precipitation data and interpolation methods commonly used in the literature to generate reference precipitation. First, we compare the errors of each RP relative to the true precipitation, derived from a super-high-density weather observation network; we then use these four RP types to evaluate the near-real-time GPM precipitation retrievals and examine the differences in evaluation results across the

various RP types. The remainder of this paper is structured as follows: Section 2 introduces the study area and the data sources, Section 3 describes the methodology employed in this study, Section 4 presents the analysis and discussion of the results, and Section 5 summarizes the main conclusions.

2. Study Area and Data

The study area for this research is Nanjing City (118.2°E–119.4°E, 31.0°N–32.8°N), located in eastern China, as illustrated in Figure 1. The terrain in this area primarily consists of plains and low hills. In this region, the topography generally has a minimal impact on precipitation, which helps to reduce terrain-related interference in the evaluation of quantitative precipitation. Nanjing experiences a humid subtropical climate influenced by the boreal monsoon, with an annual average rainfall from 800 to 1200 mm, predominantly occurring between May to September. The types of rainfall in this region include typhoon-induced rainfall and local rainfall.

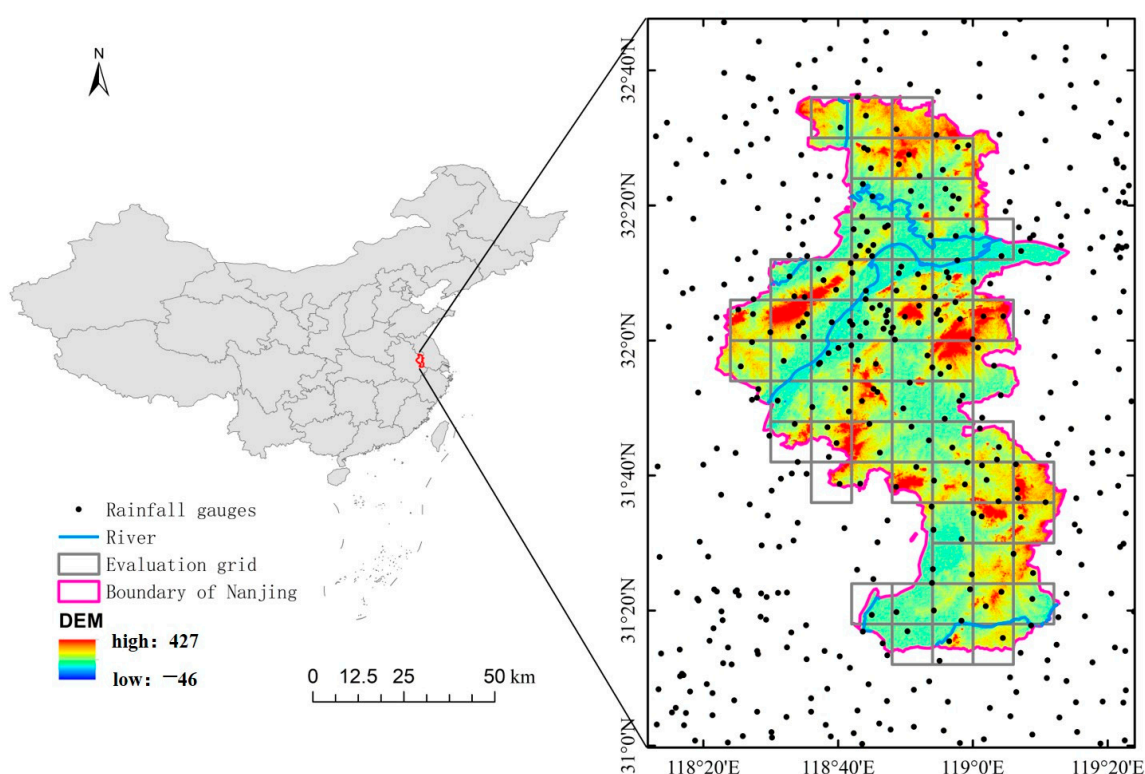


Figure 1. The geography survey of the study area.

This study focuses on an extreme rainstorm event that occurred in Nanjing City in June 2017; the rainfall broke the record for the highest single-day rainfall in Nanjing. The data utilized in this research include both ground-based rainfall gauges data and GPM precipitation products for evaluation. During this rainstorm event, 445 automatic rainfall gauges (represented as black dots on the right panel of Figure 1) deployed by the meteorological department recorded the hourly rainfall amount from 00:00 BST (Beijing Standard Time) to 23:00 BST on 10 June 2017. This dense network of rainfall gauges provides a robust foundation for generating accurate true precipitation data. Two GPM near-real-time precipitation retrieval products, namely IMERG early and IMERG late, are collected for the Nanjing region. The relevant parameters are summarized in Table 1. After processing, the hourly precipitation retrieval products that align with the station data can be obtained.

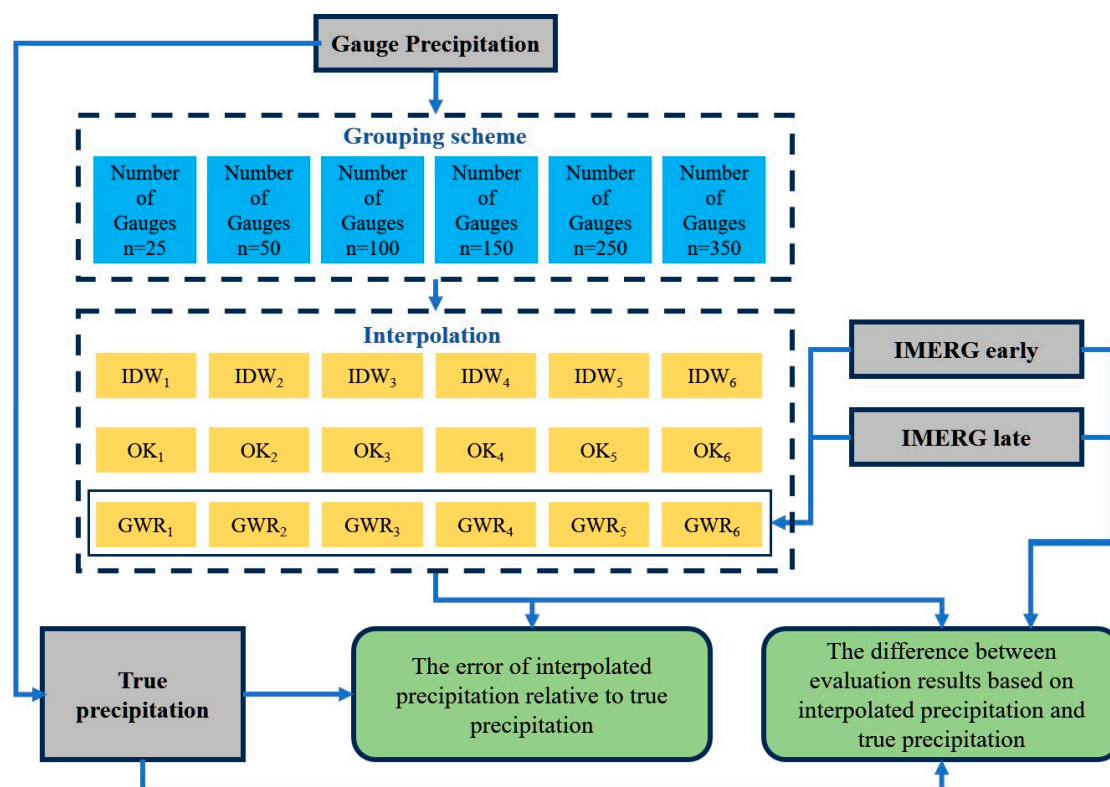
Table 1. Basic information of two GPM products.

GPM Products	Data Source	Spatial Overage	Time Resolution/h	Spatial Resolution/°
IMERG early	NASA	90°N–90°S	0.5	0.1
IMERG late	NASA	90°N–90°S	0.5	0.1

3. Methodology

3.1. Technical Scheme

Figure 2 shows the technical scheme of this study. Based on the data recorded by the 445 automatic rain gauges, various schemes are established to analyze the influences of RP on the evaluation of near-real-time satellite precipitation products. Rain gauge counts ($n = 25, 50, 100, 150, 250$, and 350) are randomly selected within the study area, with corresponding average distances between gauges of 32.6 km, 23.1 km, 16.3 km, 13.3 km, 10.3 km, and 8.7 km. To ensure the representativeness of the results, sampling for each group is repeated 100 times. To maintain spatial uniformity among gauges in each group, the 445 rainfall gauges are clustered based on geographical locations using the K-means clustering method, followed by sampling within each category.

**Figure 2.** Technical scheme.

After sampling, the inverse distance weighting (IDW), ordinary kriging (OK), and geographical weighted regression (GWR) methods are employed to interpolate the station data. Among these, the GWR method is specifically used to interpolate the difference between the observed station precipitation data and the satellite precipitation data, with satellite precipitation serving as the background field for interpolation. Depending on the chosen background fields, the interpolation results are categorized into the GWR_{early} and the GWR_{late} , which are utilized to evaluate the corresponding satellite precipitation products. The research area contains 445 ground-based rainfall gauges, resulting in a high

station density with an average of 2 gauges per grid cell. If multiple ground-based rainfall gauges are present in a grid cell, the average precipitation observed from these gauges is considered the true precipitation. In cases where no rainfall gauges are available in a grid cell, true precipitation is obtained from the eight nearest gauges using the IDW method. The results indicate that over 95% of the grid cells contain rainfall gauges, thereby ensuring the reliability of the derived true precipitation.

First, the errors of the RP obtained by different schemes relative to the true precipitation are analyzed, and a comprehensive evaluation system incorporating various indices is constructed to assess the accuracy of the RP derived from different interpolation methods and station densities. Subsequently, the near-real-time satellite precipitation product is evaluated based on the RP obtained from these schemes, and the results are compared with those derived from true precipitation.

3.2. Evaluation Indices

Quantitative indices, classification indices, volumetric categorical indices, and error decomposition indices are comprehensively used to evaluate the accuracy of near-real-time satellite precipitation products. Specifically, the quantitative indices include Correlation Coefficient (CC), Relative Bias (RB), and Root Mean Squared Error (RMSE). The Probability of Detection (POD), False Alarm Rate (FAR), and Critical Success Index (CSI) serve as classification indices. To better reflect the characteristics of heavy precipitation events, the volumetric categorical indices proposed by AghaKouchak et al. [23] are utilized in this study, namely the Volumetric Hit Index (VHI), Volumetric False Alarm Ratio (VFAR), and Volumetric Critical Success Index (VCSI), as shown in Equations (1)–(3). The values of the three volumetric categorical indices range from 0 to 1.

$$VHI = \frac{\sum_{i=1}^n (S_i | S_i > T \ \& \ G_i > T)}{\sum_{i=1}^n (S_i | S_i > T \ \& \ G_i > T) + \sum_{i=1}^n (G_i | S_i \leq T \ \& \ G_i > T)} \quad (1)$$

$$VFAR = \frac{\sum_{i=1}^n (S_i | S_i > T \ \& \ G_i \leq T)}{\sum_{i=1}^n (S_i | S_i > T \ \& \ G_i > T) + \sum_{i=1}^n (S_i | S_i > T \ \& \ G_i \leq T)} \quad (2)$$

$$VCSI = \frac{\sum_{i=1}^n (S_i | S_i > T \ \& \ G_i > T)}{\sum_{i=1}^n [S_i | (S_i > T \ \& \ G_i > T) + G_i | S_i \leq T \ \& \ G_i > T + S_i | (S_i > T \ \& \ G_i \leq T)]} \quad (3)$$

where n denotes the number of samples, S_i indicates the GPM precipitation products, and G_i indicates the ground-based RP. T refers to the 24 h or 1 h rainfall threshold for precipitation events, set at 0.1 mm.

In 1981, Willmott [24] suggested that the mean square error (MSE) of precipitation can be divided into systematic error (SE) and random error (RE) as follows:

$$MSE = SE + RE \quad (4)$$

where MSE is the total mean square error, SE is the systematic error, and RE is the random error. This equation [25] can be expressed as follows:

$$\frac{1}{n} \sum_{i=1}^n (S_i - G_i)^2 = \frac{1}{n} \sum_{i=1}^n (G_i^* - G_i)^2 + \frac{1}{n} \sum_{i=1}^n (S_i - G_i^*)^2 \quad (5)$$

$$G^* = a \times G + b \quad (6)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (S_i - G_i)^2 \quad (7)$$

$$SE = \frac{1}{n} \sum_{i=1}^n (G_i^* - G_i)^2 \quad (8)$$

$$RE = \frac{1}{n} \sum_{i=1}^n (S_i - G_i^*)^2 \quad (9)$$

$$SEP = 100\% \times \sum_{i=1}^n (G_i^* - G_i)^2 / \sum_{i=1}^n (S_i - G_i)^2 \quad (10)$$

$$REP = 100\% \times \sum_{i=1}^n (S_i - G_i^*)^2 / \sum_{i=1}^n (S_i - G_i)^2 \quad (11)$$

where S is the GPM precipitation, G^* is the regressed GPM precipitation, and G is the reference precipitation. G^* is defined using a linearity error model, where a is the slope, and b is the intercept. SEP and REP are systematic and random error percentage, respectively.

To facilitate comparisons among indices of different magnitudes, each indicator is normalized, transforming the value range of each indicator into $[0, 1]$. The worst result corresponds to 0, while the best corresponds to 1. For an evaluation index $X = (x_1, x_2, \dots, x_n)$, if it is a positive index, the worst result is initially set to 0, followed by normalization. If it is a negative index, $\text{MAX}(X^T) - X^T$ is used to convert this index to a positive form. Subsequently, the worst result of the index sequence is set to 0, and the sequence is then normalized. For an intermediate optimal index, the sequence is first converted to negative and then to positive before normalization. It should be noted that the normalized index sequences after processing inevitably include values of 0 and 1, which are dimensionless. In order to measure the overall situation of indicators, we use the sum of CI to represent the sum of various indicators, and the optimal value is 12. Table 2 shows the index normalization formula used in this study.

Table 2. Index normalization formula used in this study.

Type	Index Normalization Formula	Index
Positive index	$CI = \frac{X^T - \text{Min}(X^T)}{\text{Max}(X^T) - \text{Min}(X^T)}$	CC, POD, CSI, VHI, VCSI
Negative index	$CI = \frac{\text{Max}(X^T) - X^T}{\text{Max}(X^T) - \text{Min}(X^T)}$	RMSE, FAR, VFAR, SE, RE, SEP
Intermediate optimal index	$CI = \frac{\text{Max}(X^T - \text{Best}(X^T)) - X^T - \text{Best}(X^T) }{\text{Max}(X^T - \text{Best}(X^T)) - \text{Min}(X^T - \text{Best}(X^T))}$	RB

4. Results and Discussion

4.1. Errors of Reference Precipitation

4.1.1. Area-Averaged MAEs of RP

Figure 3 presents the boxplots of the area-averaged mean absolute errors (MAEs) of precipitation obtained by different schemes in relation to the true precipitation across experimental groups. Table 3 presents the median and mean values of the area-averaged MAE results. The results indicate that as the number of sampling gauges increases, the area-averaged MAEs derived from the IDW, $\text{GWR}_{\text{early}}$, and GWR_{late} methods gradually decrease. Specifically, when the average distance d between gauges is 8.7 km, the MAEs stabilize below 0.1 mm. In contrast, for the reference precipitation obtained through the OK method, the mean and median MAEs are 1.2 mm and 1.1 mm, respectively, at an average distance of 32.6 km. However, when the average distance d is reduced to 8.7 km, the mean and median values of the MAEs rise to 2.5 mm and 2.1 mm, respectively. This indicates that as the average distance decreases, the errors increase instead. Moreover, the increasing

trend is not always the case; i.e., the errors paradoxically increase. Moreover, this increasing trend is not consistent; the errors reach a minimum at a specific intermediate station density. This pattern is evident in both the mean and median values. When the station density is low ($n = 25$), the errors produced by the three interpolation methods are similar. Conversely, at high station densities, the errors from the IDW and GWR methods approach zero, while those from the OK method remain significant. Overall, when station density is low, the choice of interpolation methods has a minimal impact on the accuracy of the RP. However, as the station density increases, the errors associated with the OK interpolation method do not decrease but increase. Therefore, the OK method is unsuitable as an alternative interpolation technique for generating the RP.

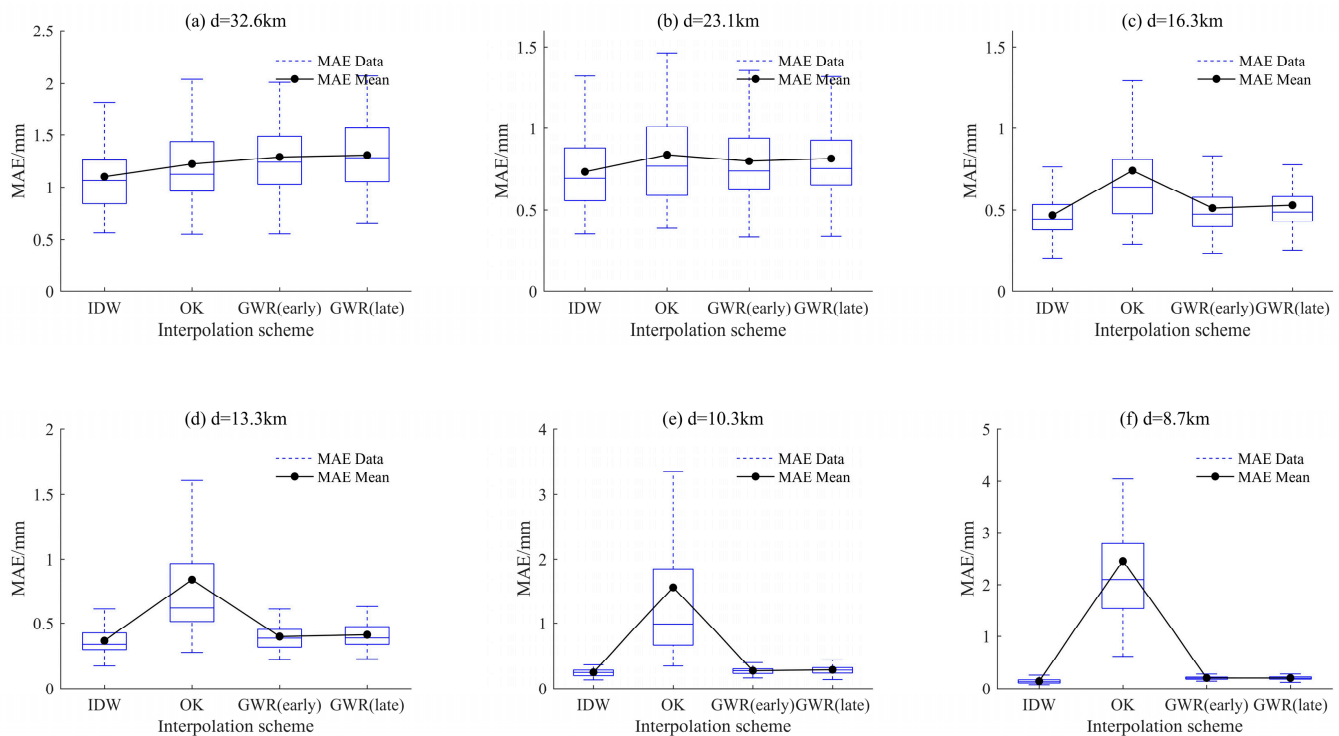


Figure 3. Boxplots of the area-averaged *MAE* of RP.

Table 3. Median and mean values of the area-averaged *MAEs* of reference precipitation.

Distance	Mean				Median			
	IDW	OK	GWR _{early}	GWR _{late}	IDW	OK	GWR _{early}	GWR _{late}
d = 32.6 km	1.1	1.2	1.3	1.3	1.1	1.1	1.2	1.3
d = 23.1 km	0.7	0.8	0.8	0.8	0.7	0.8	0.7	0.8
d = 16.3 km	0.5	0.7	0.5	0.5	0.4	0.6	0.5	0.5
d = 13.3 km	0.4	0.8	0.4	0.4	0.3	0.6	0.4	0.4
d = 10.3 km	0.3	1.6	0.3	0.3	0.2	1.0	0.3	0.3
d = 8.7 km	0.1	2.5	0.2	0.2	0.1	2.1	0.2	0.2

4.1.2. Spatial Distribution *MAEs* of RP

Based on the 100 experiments conducted for each scheme, the median *MAE* (MAE_{Median}) is selected to represent the overall performance of each scheme. Subsequently, the 24 h accumulated precipitation is compared with the true precipitation. Figure 4 illustrates the distributions of the true precipitation and the accumulated precipitation for each scheme when $MAE = MAE_{Median}$.

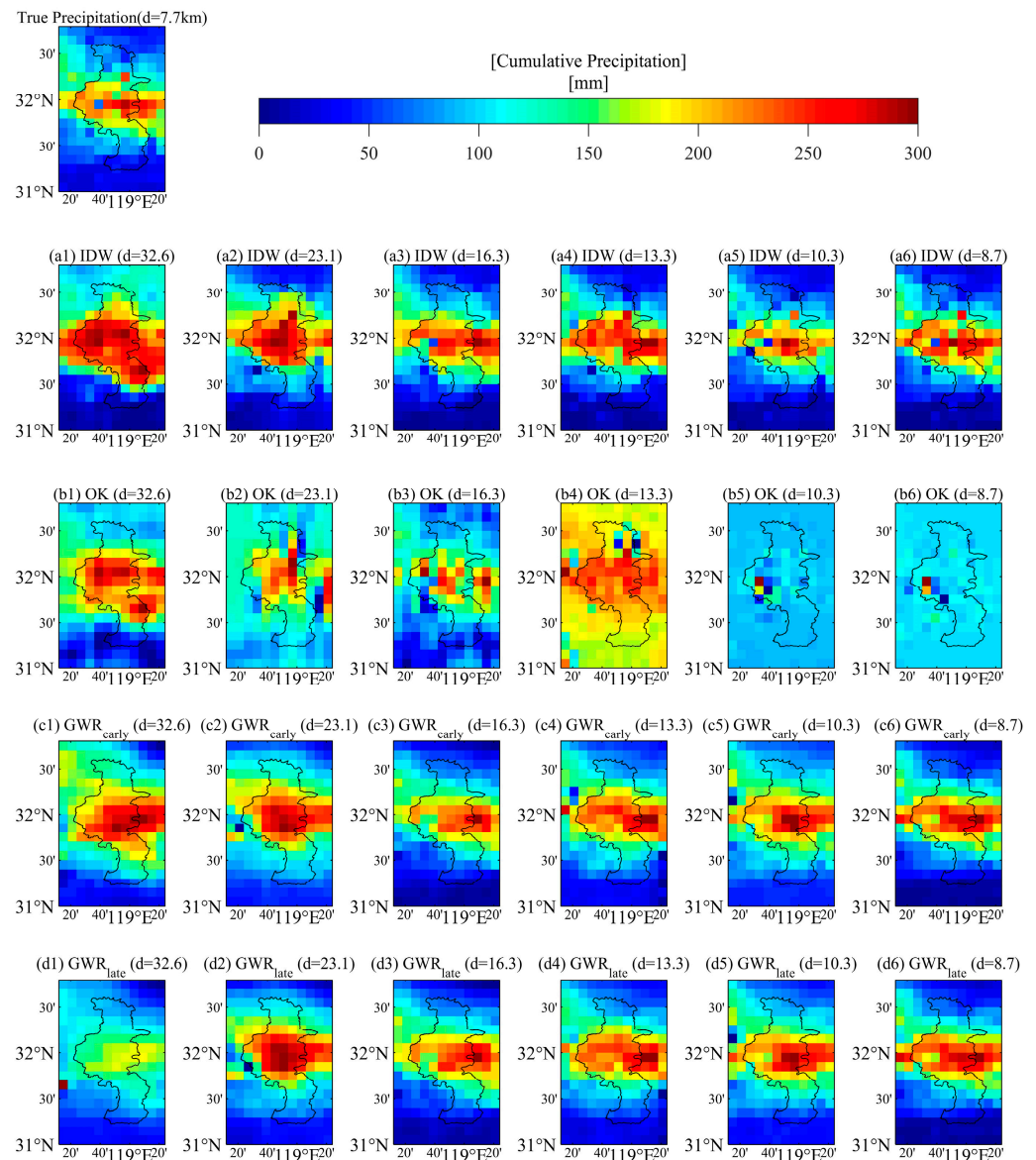


Figure 4. Spatial distribution of the accumulated precipitation when $MAE = MAE_{Median}$.

The results show that the precipitation center of the true precipitation is concentrated around 32°N, forming a strip-like pattern. As the number of gauges increases, the spatial distributions of precipitation estimated by the IDW, GWR_{early}, and GWR_{late} methods become progressively closer to the true precipitation. At low station density, the precipitation center estimated by the IDW method is noticeably larger, extending from 32°N to 31°30'N. However, the overall precipitation amount remains consistent with the true precipitation. The precipitation centers by the GWR_{early} and GWR_{late} are consistently aligned with the true precipitation center, though the total precipitation amount is significantly underestimated, particularly in the case of GWR_{late}. When station density is low, the spatial distribution of precipitation generated by the OK method is similar to that of the IDW method. While the total precipitation amount matches the true precipitation, the precipitation center is overestimated. Unlike the other interpolation methods, the spatial distribution of precipitation obtained by the OK method does not converge towards the true precipitation as the number of gauges increases. On the contrary, the deviation increases, resulting in larger discrepancies in both precipitation magnitude and center location. This finding aligns with the MAE analysis results.

4.1.3. Evaluation Indices of Errors

Figure 5 presents the boxplots of the evaluation indices (quantitative indices, classification indices, volumetric categorical indices, and error decomposition indices) of the precipitation obtained by interpolation relative to the true precipitation. Tables 4 and 5 summarize the median and mean values of these evaluation indices, respectively.

Among the quantitative indices, the CC values for precipitation interpolated by the IDW, GWR_{early}, and GWR_{late} methods increase as the mean distance between gauges decreases. When $d = 8.7$ km, the CC values exceed 0.90 for all methods. The IDW method consistently outperforms the GWR methods, although the differences diminish as the average station distance decreases. The CC value for precipitation obtained using the OK method increases initially but then declines as the station density increases, reaching a peak at $d = 16.3$ km. Unlike the CC values, the RB values for IDW, OK, and GWR methods are similar when the average station distance is large. With the decrease in the average station distance, the RB values for the OK method range between 5% and 25% when $d = 32.6$ km. The pattern of the RMSE results is similar to that of the CC results, further confirming that the IDW method has a clear advantage in precipitation interpolation. Overall, the results from the quantitative indices suggest that the OK method is unsuitable at medium to low station distances.

Regarding the classification indices, the POD values highlight the advantage of the IDW method, with the lower quartile exceeding 0.96 when $d = 32.6$ km and above 0.98 at $d = 8.7$ km. Although the GWR methods perform slightly worse than IDW, their lower quartile still exceeds 0.93 when $d = 32.6$ km. Moreover, the GWR_{early} and GWR_{late} have similar POD values at lower station densities. But as the station density increases, the POD values of the GWR_{early} are slightly larger than those of the GWR_{late}. The POD values for the OK method decrease as the station density increases. In contrast, the FAR values for GWR_{early}, GWR_{late}, and OK are consistently below 0.15, while those for IDW remain below 0.10. The FAR values of the GWR methods decline with increasing station density, with GWR_{early} showing lower FAR values than GWR_{late}. For the OK method, the FAR values decrease initially but then increase, with the lowest median FAR observed at $d = 10.3$ km, although it is still 0.10. In contrast, the IDW method's FAR median is only 0.02. The results from the CSI closely mirror those from the POD, reinforcing IDW's superiority in interpolating precipitation, with the CSI median consistently above 0.90.

The volumetric categorical indices, which weight precipitation values and focus more on heavy precipitation events, yield results similar to those of the classification indices. Thus, a detailed analysis of these results is omitted.

The error decomposition indices decompose the total error of the reference precipitation relative to the true values into systematic error and random error (RE) and calculate the proportion of systematic errors. Generally, a smaller proportion of systematic errors is preferable [24,25]. The RE values for the IDW and GWR methods gradually decrease as the station density increases, while the RE values for the OK method remain stable at larger station distances but increase sharply when the station distance decreases. There is little difference between the RE values of GWR_{early} and GWR_{late}, with IDW showing slightly lower RE values than GWR. The pattern of RE variation for the IDW and GWR methods aligns with that of the other index as the station density increases. At larger station distances, the RE values for IDW are comparable to those for OK, and the RE values for GWR_{early} are similar to those for GWR_{late}. For the OK method, the RE median decreases initially but then increases as the station density rises, reaching a minimum at $d = 16.3$ km. Regarding the SEP, the median SEP for the OK method is consistently higher than that for the IDW and GWR methods, except at $d = 8.7$ km, where it falls below that of IDW. The SEP values for IDW fluctuate, with the median consistently around 40%. The SEP values

for GWR_{early} and GWR_{late} decrease steadily as the station density increases, with minimal differences between the two methods.

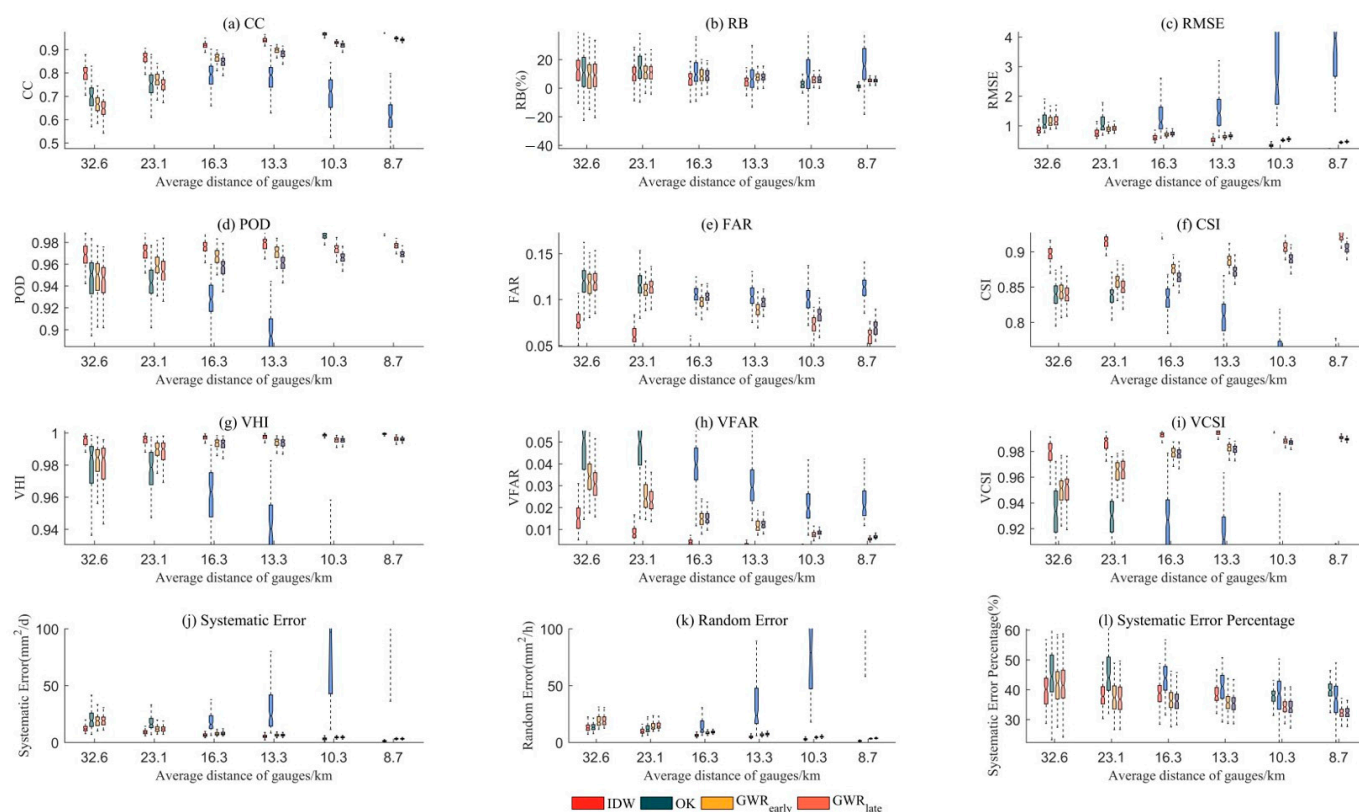


Figure 5. Boxplots of the evaluation index of interpolated precipitation relative to the true precipitation.

Table 4. Median values of the evaluation index of interpolated precipitation relative to the true precipitation.

Distance	Interpolation Scheme	CC	RB	RMSE	POD	FAR	CSI	VHI	VHAR	VCSI	SE	RE	SEP
d = 32.6 km	IDW	0.80	13.05	0.83	0.97	0.07	0.90	1.00	0.02	0.98	11.83	12.78	40.10
	OK	0.70	10.85	1.04	0.95	0.12	0.84	0.99	0.05	0.93	19.10	13.28	44.45
	GWR_{early}	0.67	9.50	1.09	0.95	0.12	0.84	0.98	0.03	0.95	19.01	17.80	42.23
	GWR_{late}	0.65	9.00	1.13	0.95	0.12	0.84	0.98	0.03	0.95	19.04	18.53	41.45
d = 23.1 km	IDW	0.87	9.75	0.72	0.97	0.06	0.92	1.00	0.01	0.99	8.95	9.36	37.73
	OK	0.75	14.11	0.98	0.94	0.12	0.84	0.98	0.05	0.93	15.93	11.21	43.99
	GWR_{early}	0.77	11.11	0.88	0.96	0.11	0.86	0.99	0.02	0.97	11.72	13.20	37.31
	GWR_{late}	0.75	11.23	0.91	0.95	0.11	0.85	0.99	0.02	0.97	11.79	14.06	36.78
d = 16.3 km	IDW	0.92	6.67	0.57	0.98	0.04	0.94	1.00	0.00	0.99	6.21	6.06	38.83
	OK	0.79	9.53	1.12	0.93	0.12	0.84	0.96	0.01	0.93	15.84	11.40	43.95
	GWR_{early}	0.87	8.23	0.69	0.97	0.10	0.88	0.99	0.01	0.98	7.39	8.11	36.26
	GWR_{late}	0.85	8.55	0.73	0.96	0.10	0.86	0.99	0.01	0.98	7.55	9.01	36.19
d = 13.3 km	IDW	0.94	4.98	0.50	0.98	0.03	0.95	1.00	0.00	0.99	5.22	4.84	37.91
	OK	0.79	6.27	1.42	0.89	0.13	0.81	0.94	0.01	0.91	23.35	23.61	41.06
	GWR_{early}	0.90	7.65	0.61	0.97	0.09	0.89	0.99	0.01	0.98	6.15	6.52	35.63
	GWR_{late}	0.88	7.52	0.65	0.96	0.10	0.87	0.99	0.01	0.98	6.31	7.27	35.45
d = 10.3 km	IDW	0.97	2.75	0.33	0.99	0.02	0.97	1.00	0.00	1.00	3.17	2.69	38.09
	OK	0.72	8.07	2.39	0.81	0.09	0.75	0.89	0.01	0.87	97.01	78.89	38.73
	GWR_{early}	0.93	5.79	0.51	0.97	0.07	0.91	1.00	0.01	0.99	4.52	4.47	34.23
	GWR_{late}	0.92	5.69	0.55	0.97	0.08	0.89	1.00	0.01	0.99	4.59	4.99	33.80
d = 8.7 km	IDW	0.99	1.16	0.16	0.99	0.01	0.98	1.00	0.00	1.00	1.15	1.23	39.85
	OK	0.61	15.39	3.96	0.75	0.10	0.69	0.86	0.01	0.85	954.26	460.99	37.08
	GWR_{early}	0.95	5.24	0.43	0.98	0.06	0.92	1.00	0.01	0.99	3.18	3.28	32.19
	GWR_{late}	0.94	5.17	0.46	0.97	0.07	0.90	1.00	0.01	0.99	3.20	3.57	32.19

Table 5. Mean values of the evaluation index of interpolated precipitation relative to the true precipitation.

Number of Sampling Gauge.	Interpolation Scheme	Index											
		CC	RB	RMSE	POD	FAR	CSI	VHI	VHAR	VCSI	SE	RE	SEP
d = 32.6 km	IDW	0.80	13.03	0.96	0.97	0.08	0.90	0.99	0.02	0.98	12.99	13.79	40.14
	OK	0.70	11.48	1.70	0.95	0.12	0.84	0.98	0.05	0.93	20.57	15.29	45.23
	GWR _{early}	0.66	9.10	1.19	0.95	0.12	0.84	0.98	0.03	0.95	21.07	34.85	41.62
	GWR _{late}	0.65	8.90	1.21	0.94	0.12	0.84	0.98	0.03	0.95	20.06	35.87	41.45
d = 23.1 km	IDW	0.86	10.24	0.83	0.97	0.06	0.91	0.99	0.01	0.99	9.67	10.49	38.38
	OK	0.75	14.61	1.24	0.94	0.12	0.84	0.97	0.05	0.93	18.41	13.15	45.57
	GWR _{early}	0.77	10.21	0.93	0.96	0.11	0.86	0.99	0.03	0.96	13.76	43.02	37.51
	GWR _{late}	0.75	10.56	0.95	0.95	0.11	0.85	0.99	0.02	0.96	13.19	33.09	37.26
d = 16.3 km	IDW	0.92	6.02	0.67	0.98	0.04	0.94	1.00	0.00	0.99	6.80	6.39	38.58
	OK	0.79	11.01	1.43	0.93	0.12	0.83	0.96	0.03	0.92	40.16	21.84	43.96
	GWR _{early}	0.86	8.54	0.74	0.97	0.10	0.88	0.99	0.02	0.98	15.17	37.74	36.58
	GWR _{late}	0.85	8.55	0.75	0.96	0.10	0.86	0.99	0.02	0.98	15.28	38.42	36.18
d = 13.3 km	IDW	0.94	4.42	0.56	0.98	0.03	0.95	1.00	0.00	0.99	5.37	5.15	38.43
	OK	0.78	6.65	5.05	0.89	0.13	0.80	0.93	0.01	0.91	108.01	49.32	40.90
	GWR _{early}	0.90	7.47	0.64	0.97	0.09	0.89	0.99	0.01	0.98	6.39	8.03	35.71
	GWR _{late}	0.88	7.47	0.73	0.96	0.10	0.87	0.99	0.01	0.98	6.44	8.20	35.42
d = 10.3 km	IDW	0.97	2.63	0.34	0.99	0.02	0.97	1.00	0.00	1.00	3.14	2.78	37.82
	OK	0.71	10.72	4.44	0.81	0.09	0.74	0.89	0.01	0.87	442.73	176.94	38.01
	GWR _{early}	0.93	5.75	0.51	0.97	0.07	0.91	1.00	0.01	0.99	4.44	4.52	34.35
	GWR _{late}	0.92	5.80	0.55	0.97	0.08	0.89	0.99	0.01	0.99	4.51	5.04	34.12
d = 8.7 km	IDW	0.99	1.36	0.16	0.99	0.01	0.98	1.00	0.00	1.00	1.20	1.24	39.89
	OK	0.61	15.94	5.60	0.76	0.08	0.69	0.86	0.01	0.85	1173.73	538.79	36.71
	GWR _{early}	0.95	5.35	0.43	0.98	0.06	0.92	1.00	0.01	0.99	3.18	3.29	32.38
	GWR _{late}	0.94	5.19	0.46	0.97	0.07	0.91	1.00	0.01	0.99	3.21	3.60	32.35

In summary, when the average station distance is large, the RP generated by the IDW method most closely approximates the true precipitation, followed by RP from the GWR method. The OK method performs significantly worse than both IDW and GWR methods in generating reference precipitation. As the average station distance decreases, the RP produced by the GWR and IDW methods becomes more aligned with the true precipitation, and the performance of the GWR method approaches that of IDW. It is noteworthy that the RP from the GWR method exhibits greater stability compared to the IDW and OK methods.

4.1.4. Comprehensive Evaluation Based on Multiple Indices

Based on the comprehensive multi-index evaluation system and the results from multiple experiments conducted within each group, the median and mean sequences are normalized and calculated by using the formula in Figure 6. The figure illustrates the comprehensive evaluation results. Among the three types of RP, the one derived from the IDW method is closest to the true precipitation, followed by the GWR method, while the OK method performs the worst.

With the decreases in the average station distance, the total composite index (CI) for the RP generated by both the IDW and GWR methods gradually increases, indicating that the RP increasingly approximates the true precipitation. Simultaneously, the gap between the RPs from the IDW and GWR methods narrows. The total CI for the OK method initially rises but then declines, peaking at $d = 16.3$ km. At larger average distances between gauges, the total CI for the GWR and OK methods are comparable, with the RP from the GWR method outperforming that from the OK method. However, at smaller average station distances, the total CI for the GWR and IDW methods is similar, with GWR_{early} and GWR_{late} showing little difference.

Based on the preceding analyses, it is evident that, in practice, a greater number of ground-based rainfall gauges leads to more accurate RPs. However, due to limitations in manpower and resources, the number of ground-based rainfall gauges is constrained. Given a reference medium station density as a basis for RP selection and evaluation, it is recommended to utilize the IDW and GWR methods to generate the RP. When $d < 20$ km, the total CI approaches 10, and the increase in total CI due to the addition of more gauges becomes negligible. Therefore, an average station distance close to 20 km is considered optimal for generating the RP.

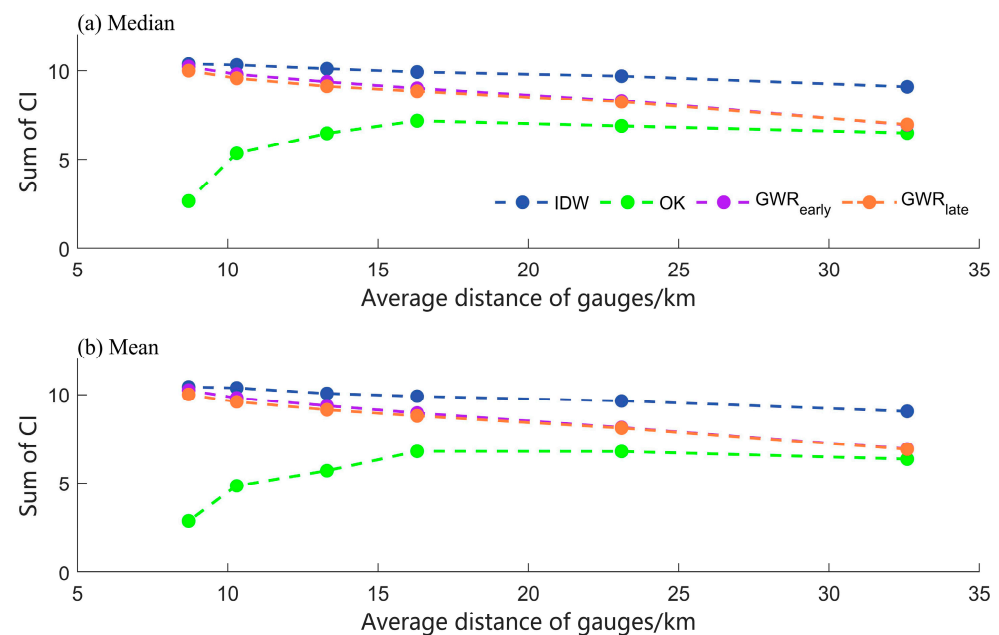


Figure 6. Comprehensive evaluation results of the median and mean values of the total composite index in multi-group experiments.

4.2. Effects of RP on the Evaluation of GPM Products

4.2.1. Comparison of Evaluation Results Based on Different RPs

In this section, we analyze the differences in the results of the true precipitation and RP for evaluating the GPM near-real-time precipitation products. Table 6 presents the median evaluation results for the GPM near-real-time precipitation products. In terms of evaluating the IMERG early and IMERG late products, the 12 indices derived from precipitation interpolated by the IDW and GWR methods gradually approach those of the near-true precipitation as the number of gauges increases. In comparison, each indicator for the IDW-based interpolated precipitation is closer to the evaluation results of the near-true precipitation. When the station count reaches 350 ($n = 350$), the difference in evaluation results between the IDW-based interpolated precipitation and the true precipitation is less than 1%, whereas the difference remains around 5% for the GWR-based interpolated precipitation. When the station number is between 50 and 100, the evaluation results for the OK-based interpolated precipitation are generally closer to those for the true precipitation. Some evaluation indices, such as the *CC*, *RB*, and *POD*, indicate that the GWR-based interpolated precipitation is close to the true precipitation, but it is clearly inferior to the IDW-based interpolated precipitation. In summary, the selection of RP has a significant impact on the evaluation of satellite precipitation products.

Table 6. Median values of the results of each RP type and the true precipitation for evaluating GPM near-real-time precipitation products.

Evaluation of NRT Satellite Precipitation	Interpolation Scheme	Distance	Index											
			<i>CC</i>	<i>RB</i>	<i>RMSE</i>	<i>POD</i>	<i>FAR</i>	<i>CSI</i>	<i>VHI</i>	<i>VHAR</i>	<i>VCSI</i>	<i>SE</i>	<i>RE</i>	<i>SEP</i>
IMERG early VS	IDW	d = 32.6 km	0.50	−44.30	2.54	0.82	0.06	0.78	0.91	0.02	0.89	39.59	12.36	65.15
		d = 23.1 km	0.49	−42.05	2.66	0.83	0.07	0.78	0.92	0.02	0.90	46.54	12.64	65.35
		d = 16.3 km	0.47	−39.58	2.67	0.83	0.08	0.78	0.92	0.03	0.90	49.04	13.07	65.11
		d = 13.3 km	0.46	−38.39	2.71	0.83	0.09	0.77	0.92	0.03	0.90	50.72	13.24	65.67
		d = 10.3 km	0.45	−38.62	2.76	0.83	0.09	0.78	0.93	0.03	0.91	53.72	13.39	65.77
		d = 8.7 km	0.45	−38.89	2.78	0.84	0.09	0.78	0.93	0.03	0.91	54.44	13.39	65.75
	OK	d = 32.6 km	0.45	−43.46	2.42	0.80	0.06	0.76	0.90	0.03	0.88	34.45	13.04	61.47
		d = 23.1 km	0.45	−42.76	2.62	0.81	0.06	0.77	0.90	0.03	0.87	43.31	13.14	63.39
		d = 16.3 km	0.43	−41.83	2.80	0.81	0.08	0.75	0.91	0.05	0.87	56.49	13.60	65.55
		d = 13.3 km	0.39	−42.11	3.15	0.81	0.12	0.73	0.91	0.09	0.84	88.67	14.21	68.89
		d = 10.3 km	0.32	−47.24	4.29	0.81	0.21	0.67	0.89	0.16	0.76	225.03	14.81	75.87
		d = 8.7 km	0.28	−54.78	7.75	0.80	0.27	0.62	0.86	0.23	0.69	1547.46	15.34	81.31
	GWR _{early}	d = 32.6 km	0.69	−45.32	1.96	0.80	0.06	0.76	0.91	0.02	0.89	23.59	8.10	63.93
		d = 23.1 km	0.66	−45.79	2.15	0.80	0.06	0.77	0.92	0.02	0.90	31.27	9.02	64.75
		d = 16.3 km	0.60	−44.45	2.32	0.81	0.06	0.77	0.92	0.02	0.90	36.24	10.41	64.15
		d = 13.3 km	0.57	−43.77	2.37	0.81	0.06	0.77	0.92	0.03	0.90	38.68	10.99	63.73
		d = 10.3 km	0.54	−43.10	2.46	0.81	0.07	0.77	0.92	0.03	0.90	41.81	11.63	64.31
		d = 8.7 km	0.53	−42.80	2.51	0.82	0.07	0.78	0.93	0.03	0.90	43.64	11.98	64.90
	True Precipitation		0.45	−38.60	2.78	0.84	0.09	0.78	0.93	0.03	0.91	53.70	13.33	65.24
IMERG late VS	IDW	d = 32.6 km	0.47	−33.37	2.10	0.82	0.09	0.76	0.92	0.04	0.89	36.70	14.07	60.76
		d = 23.1 km	0.45	−30.95	2.22	0.83	0.10	0.76	0.93	0.04	0.89	43.19	14.54	61.69
		d = 16.3 km	0.43	−27.77	2.25	0.83	0.11	0.76	0.93	0.05	0.89	45.54	14.96	61.92
		d = 13.3 km	0.42	−26.58	2.28	0.83	0.12	0.75	0.93	0.05	0.89	48.24	15.21	62.69
		d = 10.3 km	0.41	−27.11	2.32	0.83	0.12	0.76	0.93	0.05	0.89	50.92	15.39	62.77
		d = 8.7 km	0.41	−27.28	2.34	0.84	0.12	0.76	0.94	0.05	0.89	51.97	15.48	62.96
	OK	d = 32.6 km	0.42	−32.95	2.01	0.81	0.07	0.76	0.91	0.04	0.88	31.63	14.85	57.14
		d = 23.1 km	0.40	−32.18	2.17	0.82	0.08	0.76	0.91	0.05	0.87	40.26	15.11	59.68
		d = 16.3 km	0.38	−31.46	2.34	0.82	0.10	0.75	0.91	0.06	0.86	54.12	15.73	62.04
		d = 13.3 km	0.35	−31.41	2.62	0.82	0.14	0.73	0.91	0.10	0.83	85.43	16.25	66.15
		d = 10.3 km	0.29	−36.66	3.60	0.82	0.23	0.66	0.90	0.18	0.75	221.57	16.72	73.87
		d = 8.7 km	0.26	−46.35	6.50	0.81	0.28	0.61	0.87	0.24	0.68	1544.08	17.23	79.32
	GWR _{late}	d = 32.6 km	0.67	−35.36	1.59	0.81	0.09	0.75	0.92	0.04	0.89	20.67	9.35	55.33
		d = 23.1 km	0.63	−35.89	1.77	0.81	0.09	0.75	0.93	0.04	0.89	28.82	10.57	56.56
		d = 16.3 km	0.57	−34.04	1.90	0.81	0.10	0.74	0.93	0.04	0.89	33.32	12.03	57.49
		d = 13.3 km	0.54	−33.01	1.97	0.81	0.10	0.74	0.93	0.05	0.89	35.18	12.82	57.92
		d = 10.3 km	0.52	−32.15	2.04	0.82	0.10	0.75	0.93	0.05	0.89	39.07	13.34	59.20
		d = 8.7 km	0.50	−31.52	2.09	0.82	0.10	0.76	0.93	0.05	0.89	40.90	13.82	60.28
	True Precipitation		0.41	−26.96	2.34	0.84	0.12	0.76	0.94	0.05	0.89	51.26	15.47	62.45

To provide a more intuitive comparison, the same type of indices are normalized. Figure 7 illustrates the distribution of the normalized evaluation indices. The results indicate that the 12 indices for the IDW-based and GWR-based RP products increase with the station number, and the evaluation results gradually approach those of the near-true precipitation. Specifically, the *CC*, *RB*, *POD*, *VHI*, *RE*, and *SEP* values demonstrate that the IDW-based RP has clear advantages over the GWR-based RP. However, for the *VFAR*, *VCSI*, and *SE* (systematic error), the IDW-based RP and the GWR-based RP are similar. For the remaining indices, the IDW-based RP slightly outperforms the GWR-based RP. Regarding the OK-based RP, optimal values for the *CC*, Root Mean Squared Error (*RMSE*), False Alarm Rate (*FAR*), Critical Success Index (*CSI*), *VFAR*, *RE*, and *SEP* appear when the number of gauges is at a medium level. The above analysis underscores the substantial impact of RP selection on the evaluation of satellite precipitation products such as the GPM.

Combined with Figure 5, it is found that the evaluation results of the IDW-based and GWR-based RPs are close to those of the true precipitation when the average station distance is below a medium level ($d < 20$ km).

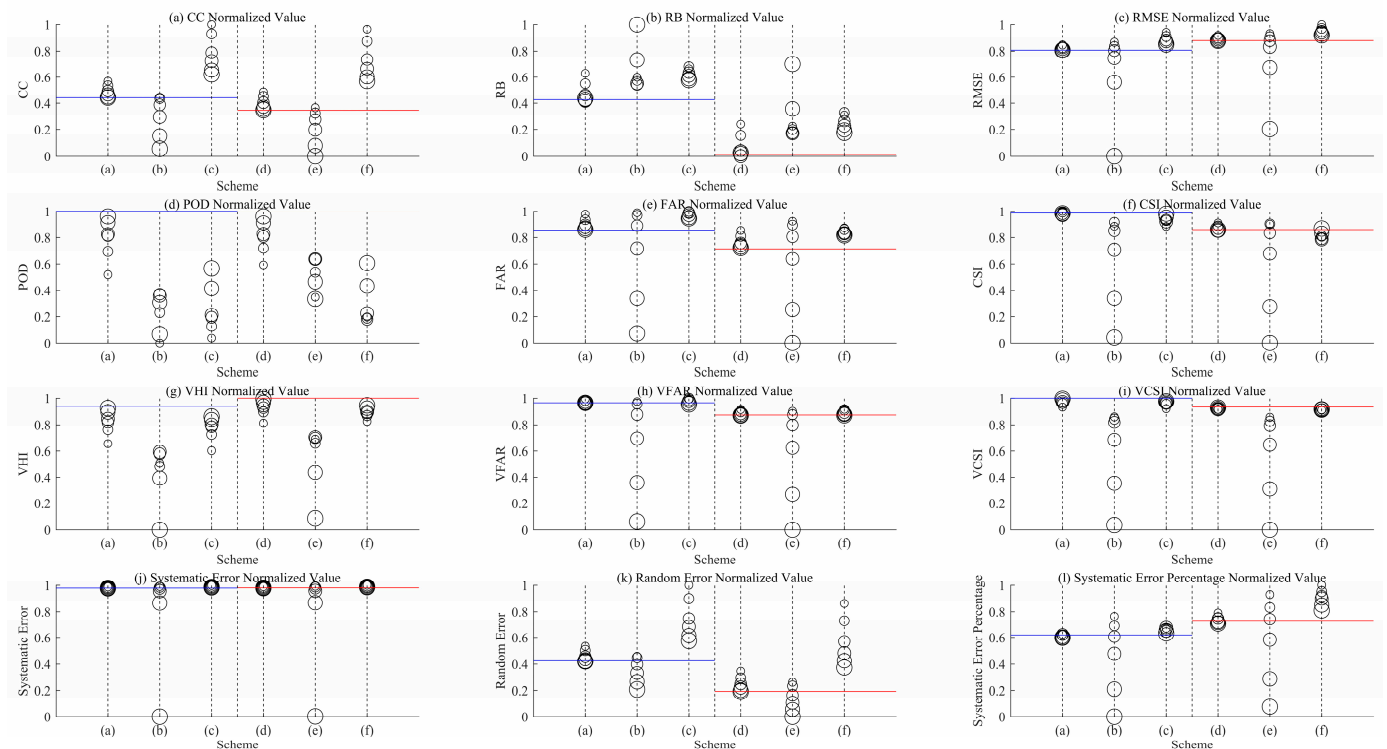


Figure 7. Normalized median distributions of the evaluation indices of the RP and true precipitation for evaluating near-real-time satellite precipitation products: (a), (b), and (c) present the RP obtained by the IDW, OK, and GWR methods to evaluate the IMERG early, and (d), (e), and (f) present the RP obtained by the IDW, OK, and GWR methods to evaluate the IMERG late. The circular symbols represent the average station distance d of 32.6, 23.1 km, 16.3 km, 13.3 km, 10.3 km, and 8.7 km. The smaller the average distance is, the larger the diameter of the circles is. The blue and red lines represent the results of the IMERG early and IMERG late products relative to the true precipitation.

4.2.2. Cause Analysis

According to the above analysis, compared with the IDW and GWR, some indices of the OK-based RP are optimal when the station number is at a particular medium level, which contradicts conventional understanding. In geostatistics, spatial properties are assumed to be homogeneous, meaning that for any point (x, y) , the mathematical expectation and variance are identical. The goal of the OK interpolation method is to determine a set of coefficients λ for a point (x, y) that results in unbiased and optimal interpolation. The semi-variance function, denoted as r , is a critical component of this method. For points with ground-based precipitation observations, it is typically calculated as $r_{ij} = (P_i - P_j)^2/2$. For unknown points, the semi-variance function values are computed based on the distances between gauges and the observations from all points with ground-based precipitation data. The relationship between station distance and the semi-variance function is usually modeled as a linear function, an exponential function, a logarithmic function, and so on. Consequently, the semi-variance function is closely linked to the distance between gauges. According to the appearance of anomalies, it can be concluded that the relationship is strongest when the station density reaches a certain value, after which the goodness of fitting decreases. The experimental results suggest that the critical average distance between gauges is between 16 km and 23 km. When the average station distance is larger than this critical value, the smaller the distance is, the better the result is. When the distance is smaller than the critical value, the interpolation results become worse and worse.

The accuracy evaluation results reveal that the accuracy of the GWR-based interpolated precipitation is inferior to that of the IDW-based interpolated precipitation. In calculating the GWR-based interpolated precipitation, the determination of the bandwidth B is crucial, and the choice of auxiliary variables directly influences the calculation of this bandwidth. Generally, introducing more auxiliary information related to the spatial distribution of precipitation allows the bandwidth to more closely approximate the true values. However, in this research, only longitude and latitude are used as the basis for calculation. Precipitation shows obvious spatial non-uniformity and is closely related to the terrain and other factors. This might be one of the reasons causing the GWR-based interpolated precipitation to be inferior to the IDW-based interpolated precipitation.

5. Conclusions

In this study, we analyze the accuracy of the RP obtained through various interpolation methods in relation to true precipitation. Subsequently, the RP is utilized to evaluate near-real-time satellite precipitation product IMERG. Additionally, a comprehensive evaluation system incorporating multiple indices is established. The main conclusions are as follows:

1. The error analysis of the RP relative to the true precipitation indicates that based on the IDW and GWR methods, the RP approaches true precipitation as the number of gauges increases. Among these methods, the RP derived from the IDW outperforms that obtained from the GWR. When the number of gauges is low, the RP from the OK method is comparable to that from the GWR method. However, as the number of gauges increases, the RP obtained via OK initially approaches true precipitation before gradually deviating from it. Results obtained from the GWR using different auxiliary variables show minimal variation.
2. The impact of different RPs on the evaluation of GPM near-real-time satellite precipitation products reveals significant discrepancies in evaluation results based on the RP used. When the average distance between gauges is below a medium level (i.e., $d < 20$ km), the evaluation results of the RPs derived from the IDW and GWR methods closely align with those of true precipitation, making them suitable for evaluating satellite precipitation products such as GPM.
3. The observed phenomenon of decreasing accuracy of the RP with an increasing number of gauges may be attributed to the overfitting of the semi-variance function to the average station distance. At a critical mean distance, the RP is closest to the true precipitation. The RP obtained via GWR is slightly less accurate than that from IDW, likely due to the limited number of auxiliary variables employed in this study. There remains potential for improvement in the RP derived from the GWR method compared to the IDW.
4. Moreover, precipitation exhibits distinct spatial non-uniformity. It should be pointed out that the critical control density is not constant, the precipitation gradually decreases with the increase in latitude, the temporal and spatial distribution of precipitation is also changing, and the representative of the station to the rainfall near the grid surface is also constantly changing, which will have a direct impact on the critical control density. This investigation focuses on Nanjing City to explore the patterns of how RP influences evaluation results, which may also display spatial non-uniformity. Further research is needed to determine whether these findings are applicable on a global scale.

Author Contributions: Conceptualization, L.W.; methodology, Y.W.; software, F.Y.; validation, Y.C.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Key R&D Program of China (2021YFC3000101), the Water Conservancy Technology Project of Jiangsu Province (202311), the National Natural Science Foundation of China (52109028, 52239008), and the Natural Science Foundation of Jiangsu Province (BK20210042).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Yilan Li and Yiceng Chen were employed by the company Wuhan Zhongwei Intelligent Survey Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Hong, Y.; Gochis, D.; Cheng, J.-T.; Hsu, K.-L.; Sorooshian, S. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *J. Hydrometeorol.* **2007**, *8*, 469–482. [\[CrossRef\]](#)
2. Joyce, R.J.; Janowiak, J.E.; Arkin, P.A.; Xie, P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* **2004**, *5*, 487–503. [\[CrossRef\]](#)
3. 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. [\[CrossRef\]](#)
4. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Oki, R.; Nakamura, K.; Iguchi, T. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 701–722. [\[CrossRef\]](#)
5. Tang, G.; Zeng, Z.; Ma, M.; Liu, R.; Wen, Y.; Hong, Y. Can near-real-time satellite precipitation products capture rainstorms and guide flood warning for the 2016 summer in South China? *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1208–1212. [\[CrossRef\]](#)
6. Prakash, S.; Mitra, A.K.; Pai, D.S.; AghaKouchak, A. From TRMM to GPM: How well can heavy rainfall be detected from space? *Adv. Water Resour.* **2016**, *88*, 1–7. [\[CrossRef\]](#)
7. Shawky, M.; Moussa, A.; Hassan, Q.K.; El-Sheimy, N. Performance assessment of sub-daily and daily precipitation estimates derived from GPM and GSMaP products over an arid environment. *Remote Sens.* **2019**, *11*, 2840. [\[CrossRef\]](#)
8. Libertino, A.; Sharma, A.; Lakshmi, V.; Claps, P. A global assessment of the timing of extreme rainfall from TRMM and GPM for improving hydrologic design. *Environ. Res. Lett.* **2016**, *11*, 054003. [\[CrossRef\]](#)
9. Sahlu, D.; Nikolopoulos, E.I.; Moges, S.A.; Anagnostou, E.N.; Hailu, D. First evaluation of the Day-1 IMERG over the upper Blue Nile basin. *J. Hydrometeorol.* **2016**, *17*, 2875–2882. [\[CrossRef\]](#)
10. He, Z.; Yang, L.; Tian, F.; Ni, G.; Hou, A.; Lu, H. Intercomparisons of rainfall estimates from TRMM and GPM multisatellite products over the Upper Mekong River Basin. *J. Hydrometeorol.* **2017**, *18*, 413–430. [\[CrossRef\]](#)
11. Hayashi, Y.; Tebakari, T.; Hashimoto, A. A Comparison Between Global Satellite Map** of Precipitation Data and High-Resolution Radar Data—A Case Study of Localized Torrential Rainfall over Japan. *J. Disaster Res.* **2021**, *16*, 786–793. [\[CrossRef\]](#)
12. Mekonnen, K.; Melesse, A.M.; Woldesenbet, T.A. Spatial evaluation of satellite-retrieved extreme rainfall rates in the Upper Awash River Basin, Ethiopia. *Atmos. Res.* **2021**, *249*, 105297. [\[CrossRef\]](#)
13. Hosseini-Moghari, S.M.; Tang, Q. Validation of GPM IMERG V05 and V06 precipitation products over Iran. *J. Hydrometeorol.* **2020**, *21*, 1011–1037. [\[CrossRef\]](#)
14. Zhou, Z.; Guo, B.; Xing, W.; Zhou, J.; Xu, F.; Xu, Y. Comprehensive evaluation of latest GPM era IMERG and GSMaP precipitation products over mainland China. *Atmos. Res.* **2020**, *246*, 105132. [\[CrossRef\]](#)
15. Tang, G.; Behrangi, A.; Long, D.; Li, C.; Hong, Y. Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *J. Hydrol.* **2018**, *559*, 294–306. [\[CrossRef\]](#)
16. Tan, M.L.; Ibrahim, A.L.; Duan, Z.; Cracknell, A.P.; Chaplot, V. Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sens.* **2015**, *7*, 1504–1528. [\[CrossRef\]](#)
17. Gadelha, A.N.; Coelho, V.H.R.; Xavier, A.C.; Barbosa, L.R.; Melo, D.C.; Xuan, Y.; Huffman, G.J.; Petersen, W.A.; Almeida, C.d.N. Grid box-level evaluation of IMERG over Brazil at various space and time scales. *Atmos. Res.* **2019**, *218*, 231–244. [\[CrossRef\]](#)
18. Nepal, B.; Shrestha, D.; Sharma, S.; Shrestha, M.S.; Aryal, D.; Shrestha, N. Assessment of GPM-Era Satellite Products (IMERG and GSMaP) ability to detect precipitation extremes over mountainous country Nepal. *Atmosphere* **2021**, *12*, 254. [\[CrossRef\]](#)
19. Hu, Q.; Yang, H.; Meng, X.; Wang, Y.; Deng, P. Satellite and gauge rainfall merging using geographically weighted regression. *Proc. Int. Assoc. Hydrol. Sci.* **2015**, *368*, 132–137. [\[CrossRef\]](#)
20. Darand, M.; Amanollahi, J.; Zandkarimi, S. Evaluation of the performance of TRMM Multi-satellite Precipitation Analysis (TMPA) estimation over Iran. *Atmos. Res.* **2017**, *190*, 121–127. [\[CrossRef\]](#)

21. Tian, B.; Chen, H.; Wang, J.; Xu, C.-Y. Accuracy assessment and error cause analysis of GPM (V06) in Xiangjiang river catchment. *Hydrol. Res.* **2021**, *52*, 1048–1065. [[CrossRef](#)]
22. Qureshi, S.; Koohpayma, J.; Firozjaei, M.K.; Kakroodi, A.A. Evaluation of seasonal, drought, and wet condition effects on performance of satellite-based precipitation data over different climatic conditions in Iran. *Remote Sens.* **2021**, *14*, 76. [[CrossRef](#)]
23. AghaKouchak, A.; Mehran, A. Extended contingency table: Performance metrics for satellite observations and climate model simulations. *Water Resour. Res.* **2013**, *49*, 7144–7149. [[CrossRef](#)]
24. Willmott, C.J. On the validation of models. *Phys. Geogr.* **1981**, *2*, 184–194. [[CrossRef](#)]
25. Habib, E.; Henschke, A.; Adler, R.F. Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmos. Res.* **2009**, *94*, 373–388. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.