remote sensing

Remote Sensing of Precipitation
Part II

Edited by
Silas Michaelides

Printed Edition of the Special Issue Published in *Remote Sensing*

MDPI

Remote Sensing of Precipitation: Part II

Remote Sensing of Precipitation: Part II

Editor

Silas Michaelides

MDPI

Editor
Silas Michaelides
Climate and Atmosphere
Research Center (CARE-C)
The Cyprus Institute
Cyprus

This is a reprint of articles from the Special Issue published online in the open access journal *Remote Sensing* (ISSN 2072-4292) (available at: https://www.mdpi.com/journal/remotesensing/special_issues/_precipitation2).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

Contents

About the Editor

Silas Michaelides Dr Silas Michaelides is an Adjunct Professor at the Climate and Atmosphere Research Center (CARE-C) of The Cyprus Institute. He has been the Director of the Department of Meteorology of Cyprus, a position he has reached having climbed all the scientific ranks of this governmental organization, for more than 40 years. He holds a PhD in Meteorology, an MSc in Agricultural Meteorology, a Master's degree in Public Sector Management and a BSc in Mathematics. He has published 125 papers in peer-reviewed scientific journals, several of which are on the remote sensing of precipitation. He has also published one book on precipitation, and he has acted as the Guest Editor of more than 30 Special Issues and Conference Proceedings. He is a Member Emeritus of the European Geosciences Union, a Member of the American Meteorological Society, a Fellow of the Royal Meteorological Society and a Member of the Hellenic Meteorological Society. He is currently the Vice Chairman of the Cyprus Remote Sensing Society.

Editorial

Editorial for Special Issue "Remote Sensing of Precipitation: Part II"

Silas Michaelides

Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, 2121 Aglantzia, Cyprus; s.michaelides@cyi.ac.cy

1. Introduction

The ongoing and intensive consideration by the scientific community of the many facets of precipitation science constitutes a broad recognition of the significance of this indispensable component of the hydrologic cycle. The interest in the state-of-the-art scientific investigations of precipitation is maintained and even rejuvenated by current developments in the field, embracing the availability of new precipitation databases, technological improvements and methodological advances.

This Special Issue comprises a collection of papers embracing a wide range of aspects of precipitation science. This volume hosts 25 papers devoted to remote sensing applications in precipitation, which include studies on satellite precipitation retrievals together with their corresponding methodologies, weather radar precipitation estimations, the understanding of cloud and precipitation microphysical properties, precipitation downscaling, droplet size distribution, the performance of precipitation forecasts by numerical weather prediction models, precipitation retrievals from global navigation satellite systems, spatiotemporal precipitation distribution and its statistical characteristics, rain gauge- and satellite-based precipitation products and their comparisons and spatiotemporal distributions of precipitation modeling.

The next section summarizes the individual articles hosted in this Special Issue in alphabetical order based on the first author's name.

2. Overview of Contributions

Aminyavari et al. [1] studied the performance of ensemble forecasts for precipitation estimates by three numerical weather prediction models within THORPEX (The Observing System Research and Predictability Experiment of the World Meteorological Organization), as well as that of the IMERG (Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (GPM)), during severe floods in Iran over the period of March and April 2019.

In the paper by D'Adderio et al. [2], precipitation estimates derived from the Italian ground radar network were used in conjunction with measurements from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) to develop an operational-oriented algorithm (RAdar INfrared Blending algorithm for Operational Weather monitoring—RAINBOW) able to provide precipitation pattern and intensity. The algorithm evaluated surface precipitation over five geographical boxes in which the study area was divided.

In their study, Eldardiry and Habib [3] assessed the robustness of a probability weighted regional spatial bootstrap approach to estimate precipitation frequencies using radar data. Using the regional spatial bootstrap technique, they investigated two main issues that impact the use of radar-based Quantitative Precipitation Estimations (QPE) in deriving precipitation frequency estimates: (1) the typically short historical records of radar-based QPEs and (2) the effect of outliers in a precipitation maxima series that could possibly cause unrealistic spatial gradients in intensity–duration–frequency relations.

Ghada et al. [4] contributed with an analysis of the relationship between Z–R (reflectivity and rain rate) parameters and weather types in Central Europe, based on a

comprehensive regional dataset of rain microstructure measurements at ten sites in the federal state of Bavaria, Germany. The authors investigated what the effect is of weather types on rain microstructures and whether there is a consistent variation in the Z–R parameters between weather types that would suggest opportunities to improve the QPE with radar-based methods.

The study by Giannaros et al. [5] presented the first attempt for introducing an assimilation of zenith tropospheric delay, derived from more than 48 stations of the Hellenic global navigation satellite systems network, into the operational Numerical Weather Prediction System of the National Observatory of Athens in Greece, which is based on the Mesoscale Weather Research and Forecasting model. The model was applied during seven high-impact precipitation events covering the dry and wet seasons in 2018.

The main objective of the study by Huang et al. [6] was to evaluate the performance of multiple Satellite Precipitation Products (SSPs) in depicting the spatiotemporal variations of summer precipitation over Taiwan, using more than 400 local rain gauges for comparison. The authors examined the competence of SSPs in studying the summer connective afternoon rainfall events, which constitute the most frequently observed weather patterns in Taiwan. The analysis mainly focused on the time periods that overlap in all the data investigated—that is, the summers of 2014–2017.

Precipitation measurements from a second-generation PARSIVEL disdrometer deployed in Beijing, Northern China, were analyzed by Ji et al. [7] in order to investigate the microphysical structure of raindrop size distribution and its implications on polarimetric radar applications. Rainfall types were classified and analyzed in the domains of the median volume diameter and the normalized intercept parameter.

Krietemeyer et al. [8] investigated the performance of the Precise Point Positioning technique for the estimation of the zenith tropospheric delays by using a recently introduced low-cost dual-frequency receiver connected to antennas of ranging quality with and without applying relative antenna calibrations. Additionally, using Level 1 data, they investigated how well the (un-)corrected single-frequency data from the dual-frequency receiver can be used for meteorological applications.

Laverde-Barajas et al. [9] focused their research on a spatiotemporal object-based bias correction method to reduce several systematic errors in storm events estimated by satellite. The method, called Spatiotemporal Contiguous Object-based Rainfall Analysis for Bias Correction, uses the main storm characteristics of the satellite and observed events detected to remove errors due to displacement in space and time and volume. This method was evaluated over the lower Mekong Basin in Thailand to correct several storm event types in IMERG during the monsoon seasons from 2014 to 2017.

Le Coz et al. [10] proposed to gauge-adjust satellite-based estimates with respect to position by using a morphing method. This approach takes both the position and the intensity of a rain event into account, and its potential was investigated with two case studies. In the first case, the rain events were synthetic, represented by elliptic shapes, while, in the second case, use was made of real data from a rainfall event occurring during the monsoon season in Southern Ghana. In the second case, the satellite-based estimate IMERG was adjusted to gauge data from the Trans-African Hydro-Meteorological Observatory (TAHMO) network.

In the framework of JAXA's Global Satellite Mapping of Precipitation (GSMaP) project, the GSMaP Near-Real-Time Precipitation Products were generated (GSMaP_Gauge_NRT), aiming to improve the accuracy of the near-real-time product of GSMaP. Lu and Yong [11] used GSMaP_Gauge_NRT to validate their performance by using gauge observations over Mainland China.

Lu et al. [12] proposed a two-step framework to improve the accuracy of satellite precipitation estimates. The first step was data merging based on the optimum interpolation, and the second step was downscaling based on geographically weighted regression. The IMERG product was used to demonstrate the effectiveness of the above two-step procedure in the Tianshan Mountains, China.

The primary purpose of the study by Ma et al. [13] was to conduct a comprehensive analysis of an extremely heavy rainfall event that hit Guangdong Province, China, from 27 August to 1 September 2018. Their analysis was based on various in situ and remote sensing observations, including rain gauges, polarimetric radars, disdrometers and reanalysis data, to gain a better understanding of the epic flood events. In particular, the study aimed to explore the potential of polarimetric radars to resolve the microphysics and quantify the precipitation.

Maghsood et al. [14] conducted a thorough validation of the IMERG product over Iran. The study focused on investigating the performance of daily and monthly IMERG products by comparing them with ground-based precipitation data at synoptic stations throughout the country during 2014–2017. The spatial and temporal performance of the IMERG was evaluated using eight statistical criteria.

Nawaz et al. [15] highlighted the performance evaluation of gauge-based and satellite-based gridded precipitation products at the annual, winter and summer monsoon scales by using a multiple statistical approach over Punjab Province, Pakistan. The results revealed that the temporal magnitude of all the gridded precipitation products was different and deviated up to 100–200 mm with the overall spatial pattern of underestimation and overestimation from north to south.

Rahman et al. [16] assessed the performance of a soil moisture-to-rain algorithm that can be used for direct precipitation estimation from in-situ and/or satellite-based soil moisture observations. The algorithm makes use of the inverted soil water balance equation. Their study covered four different climate regions of Pakistan using rain gauge observations. The assessment was carried out on a daily scale in the period 2000–2015.

Retalis et al. [17] performed a comparative analysis of the IMERG high-resolution product and Tropical Rainfall Measuring Mission (TRMM) 3B43 product. These satellite-based precipitation fields were validated against rain gauges over the island of Cyprus for the period from April 2014 to June 2018. Satellite precipitation estimates were compared with the gauge records on a monthly and an annual basis.

In an effort to elucidate the strengths and weaknesses of recently released gridded precipitation datasets, Sherifi et al. [18] conducted a comprehensive evaluation of the performances of several such datasets at daily and monthly timescales. The study was performed over Austria using a dense network of gauges of 882 stations. The evaluation was carried out based on continuous and categorical statistical metrics for the period from June 2014 to December 2015.

Sharma et al. [19] evaluated four precipitation datasets from the two satellite-based precipitation products, namely IMERG (version 06B) and GSMaP (version 07), against 388 rain gauge observations concerning their spatial and seasonal accuracy over Nepal. Their performances were analyzed for their tendencies and discrepancies depending on the different elevation ranges and relative intensities on a daily and monthly timescale from March 2014 to December 2016.

The study by Sokol et al. [20] focused on Linear Depolarization Ratio values derived from vertically pointing cloud radars and the distribution of five hydrometeor species during 38 days with thunderstorms that occurred in 2018 and 2019 in Central Europe within the vicinity of the radar used in the study. The study showed improved algorithms for de-aliasing, the derivation of vertical air velocity and the classification of hydrometeors in clouds.

The main objectives of the contribution by Sun et al. [21] were (a) to investigate the spatial distribution of precipitation changes using daily, daytime and nighttime records from 2393 weather sites across China; (b) to study the quantification of the performances of selected products in detecting precipitation trends on a sub-daily scale with different validation metrics through a comparison with gauge observations and (c) to identify the metric-based optimal products at a sub-daily scale.

In the research by Ullah et al. [22], a new downscaling methodology was developed using the Digital Elevation Model to delineate into three geospatial predictors, i.e., elevation,

longitude and latitude, in an empirical distribution-based framework (EDBF). Two different satellite-based precipitation datasets, such as the GPM-based multitemporal precipitation data for the prediction of high-resolution downscaled weighted precipitation from 0.1° to 0.05° > resolution, and the GPM and the TRMM datasets for the verification of proposed methodology were used over the humid southern region of Mainland China.

In the study by Xie et al. [23], an OTT Parsivel-2 (Kempten, Germany) Disdrometer was used to measure raindrop spectra from 10 August 2018 to 10 August 2019 at Yulin Ecohydrological Station, Shaanxi Province, China. The precipitation events obtained were classified as stratiform and convective, based on the rainfall intensity classifying process. The Drop Size Distribution characteristics of Yulin Station were obtained, and the results can help to understand the microphysical characteristics of precipitation and their impact on the mechanism of soil erosivity in the semi-arid area.

The study by Yang et al. [24] aimed to evaluate the accuracy of the latest five GPM and TRMM rainfall products across monthly, daily and hourly scales based on ground rain gauge measurements between January 2009 and December 2017 in the Shuaishui River Basin of Eastern-Central China. For the evaluation, a total of four continuous and three categorical metrics were calculated based on satellite precipitation product (SPP) estimates and historical rainfall records at 13 stations over a period of nine years from 2009 to 2017.

Zhang et al. [25] established a spatial and temporal distribution model of precipitation in Hubei Province, China from 2006 to 2014 based on the data of 75 meteorological stations. This paper applied a geographically and temporally weighted regression kriging model to precipitation and assessed the effects of timescales and a time-weighted function on precipitation interpolation.

3. Conclusions

This Special Issue aimed at enlightening and updating the scientific community involved in precipitation science to the current progress in important areas of the remote sensing of precipitation through the presentation of state-of-the-art data sources and technological advances, as well as relevant methodological approaches. This collection of papers aspires to stimulate further research in the remote sensing of precipitation.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Aminyavari, S.; Saghafian, B.; Sharifi, E. Assessment of Precipitation Estimation from the NWP Models and Satellite Products for the Spring 2019 Severe Floods in Iran. *Remote Sens.* **2019**, *11*, 2741. [CrossRef]
2. D'Adderio, L.P.; Puca, S.; Vulpiani, G.; Petracca, M.; Sanò, P.; Dietrich, S. RAINBOW: An Operational Oriented Combined IR-Algorithm. *Remote Sens.* **2020**, *12*, 2444. [CrossRef]
3. Eldardiry, H.; Habib, E. Examining the Robustness of a Spatial Bootstrap Regional Approach for Radar-Based Hourly Precipitation Frequency Analysis. *Remote Sens.* **2020**, *12*, 3767. [CrossRef]
4. Ghada, W.; Bech, J.; Estrella, N.; Hamann, A.; Menzel, A. Weather Types Affect Rain Microstructure: Implications for Estimating Rain Rate. *Remote Sens.* **2020**, *12*, 3572. [CrossRef]
5. Giannaros, C.; Kotroni, V.; Lagouvardos, K.; Giannaros, T.M.; Pikridas, C. Assessing the Impact of GNSS ZTD Data Assimilation into the Wrf Modeling System during High-Impact Rainfall Events over Greece. *Remote Sens.* **2020**, *12*, 383. [CrossRef]

6. Huang, W.R.; Liu, P.Y.; Chang, Y.H.; Liu, C.Y. Evaluation and Application of Satellite Precipitation Products in Studying the Summer Precipitation Variations over Taiwan. *Remote Sens.* **2020**, *12*, 347. [CrossRef]
7. Ji, L.; Chen, H.; Li, L.; Chen, B.; Xiao, X.; Chen, M.; Zhang, G. Raindrop Size Distributions and Rain Characteristics Observed by a PARSIVEL Disdrometer in Beijing, Northern China. *Remote Sens.* **2019**, *11*, 1479. [CrossRef]
8. Krietemeyer, A.; van der Marel, H.; van de Giesen, N.; ten Veldhuis, M.C. High Quality Zenith Tropospheric Delay Estimation Using a Low-Cost Dual-Frequency Receiver and Relative Antenna Calibration. *Remote Sens.* **2020**, *12*, 1393. [CrossRef]
9. Laverde-Barajas, M.; Corzo, G.A.; Poortinga, A.; Chishtie, F.; Meechaiya, C.; Jayasinghe, S.; Towashiraporn, P.; Markert, A.; Saah, D.; Son, L.H.; et al. St-Corabico: A Spatiotemporal Object-Based Bias Correction Method for Storm Prediction Detected by Satellite. *Remote Sens.* **2020**, *12*, 3538. [CrossRef]
10. Le Coz, C.; Heemink, A.; Verlaan, M.; ten Veldhuis, M.C.; van de Giesen, N. Correcting Position Error in Precipitation Data Using Image Morphing. *Remote Sens.* **2019**, *11*, 2557. [CrossRef]
11. Lu, D.; Yong, B. A Preliminary Assessment of the Gauge-Adjusted near-Real-Time GSMaP Precipitation Estimate over Mainland China. *Remote Sens.* **2020**, *12*, 141. [CrossRef]
12. Lu, X.; Tang, G.; Wang, X.; Liu, Y.; Wei, M.; Zhang, Y. The Development of a Two-Step Merging and Downscaling Method for Satellite Precipitation Products. *Remote Sens.* **2020**, *12*, 398. [CrossRef]
13. Ma, Y.; Chen, H.; Ni, G.; Chandrasekar, V.; Gou, Y.; Zhang, W. Microphysical and Polarimetric Radar Signatures of an Epic Flood Event in Southern China. *Remote Sens.* **2020**, *12*, 2772. [CrossRef]
14. Maghsood, F.F.; Hashemi, H.; Hosseini, S.H.; Berndtsson, R. Ground Validation of GPM IMERG Precipitation Products over Iran. *Remote Sens.* **2020**, *12*, 48. [CrossRef]
15. Nawaz, Z.; Li, X.; Chen, Y.; Nawaz, N.; Gull, R.; Elnashar, A. Spatio-Temporal Assessment of Global Precipitation Products over the Largest Agriculture Region in Pakistan. *Remote Sens.* **2020**, *12*, 3650. [CrossRef]
16. Rahman, K.U.; Shang, S.; Shahid, M.; Wen, Y. Performance Assessment of SM2RAIN-CCI and SM2RAIN-ASCAT Precipitation Products over Pakistan. *Remote Sens.* **2019**, *11*, 2040. [CrossRef]
17. Retalis, A.; Katsanos, D.; Tymvios, F.; Michaelides, S. Comparison of GPM Imerg and TRMM 3B43 Products over Cyprus. *Remote Sens.* **2020**, *12*, 3212. [CrossRef]
18. Sharifi, E.; Eitzinger, J.; Dorigo, W. Performance of the State-of-the-Art Gridded Precipitation Products over Mountainous Terrain: A Regional Study over Austria. *Remote Sens.* **2019**, *11*, 2018. [CrossRef]
19. Sharma, S.; Chen, Y.; Zhou, X.; Yang, K.; Li, X.; Niu, X.; Hu, X.; Khadka, N. Evaluation of GPM-Era Satellite Precipitation Products on the Southern Slopes of the Central Himalayas against Rain Gauge Data. *Remote Sens.* **2020**, *12*, 1836. [CrossRef]
20. Sokol, Z.; Minářová, J.; Fišer, O. Hydrometeor Distribution and Linear Depolarization Ratio in Thunderstorms. *Remote Sens.* **2020**, *12*, 2144. [CrossRef]
21. Sun, S.; Shi, W.; Zhou, S.; Chai, R.; Chen, H.; Wang, G.; Zhou, Y.; Shen, H. Capacity of Satellite-Based and Reanalysis Precipitation Products in Detecting Long-Term Trends across Mainland China. *Remote Sens.* **2020**, *12*, 2902. [CrossRef]
22. Ullah, S.; Zuo, Z.; Zhang, F.; Zheng, J.; Huang, S.; Lin, Y.; Iqbal, I.; Sun, Y.; Yang, M.; Yan, L. Gpm-Based Multitemporal Weighted Precipitation Analysis Using Gpm_imergdf Product and Aster Dem in Edbf Algorithm. *Remote Sens.* **2020**, *12*, 3162. [CrossRef]
23. Xie, Z.; Yang, H.; Lv, H.; Hu, Q. Seasonal Characteristics of Disdrometer-Observed Raindrop Size Distributions and Their Applications on Radar Calibration and Erosion Mechanism in a Semi-Arid Area of China. *Remote Sens.* **2020**, *12*, 262. [CrossRef]
24. Yang, X.; Lu, Y.; Tan, M.L.; Li, X.; Wang, G.; He, R. Nine-Year Systematic Evaluation of the GPM and TRMM Precipitation Products in the Shuaishui River Basin in East-Central China. *Remote Sens.* **2020**, *12*, 1042. [CrossRef]
25. Zhang, W.; Liu, D.; Zheng, S.; Liu, S.; Loáiciga, H.A.; Li, W. Regional Precipitation Model Based on Geographically and Temporally Weighted Regression Kriging. *Remote Sens.* **2020**, *12*, 2547. [CrossRef]

MDPI

Article

Examining the Robustness of a Spatial Bootstrap Regional Approach for Radar-Based Hourly Precipitation Frequency Analysis

Hisham Eldardiry [1,†] and Emad Habib [2,*]

[1] Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, LA 70503, USA; dardiry@uw.edu
[2] Department of Civil Engineering and Watershed Flood Center, University of Louisiana at Lafayette, Lafayette, LA 70503, USA
* Correspondence: emad.habib@louisiana.edu; Tel.: +1-337-482-6513
† Currently address: Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA.

Abstract: Radar-based Quantitative Precipitation Estimates (QPE) provide rainfall products with high temporal and spatial resolutions as opposed to sparse observations from rain gauges. Radar-based QPE's have been widely used in many hydrological and meteorological applications; however, using these high-resolution products in the development of Precipitation Frequency Estimates (PFE) is impeded by their typically short-record availability. The current study evaluates the robustness of a spatial bootstrap regional approach, in comparison to a pixel-based (i.e., at site) approach, to derive PFEs using hourly radar-based multi-sensor precipitation estimation (MPE) product over the state of Louisiana in the US. The spatial bootstrap sampling technique augments the local pixel sample by incorporating rainfall data from surrounding pixels with decreasing importance when distance increases. We modeled extreme hourly rainfall data based on annual maximum series (AMS) using the generalized extreme value statistical distribution. The results showed a reduction in the uncertainty bounds of the PFEs when using the regional spatial bootstrap approach compared to the pixel-based estimation, with an average reduction of 10% and 2% in the 2- and 5-year return periods, respectively. Using gauge-based PFE's as a reference, the spatial bootstrap regional approach outperforms the pixel-based approach in terms of robustness to outliers identified in the radar-based AMS of some pixels. However, the systematic bias inherent to radar-based QPE especially for extreme rainfall cases, appear to cause considerable underestimation in PFEs in both the pixel-based and the regional approaches.

Keywords: rainfall; radar; extreme precipitation; spatial bootstrap; Louisiana; annual maxima

1. Introduction

Rainfall plays a critical role in the earth's water and energy cycle over a wide range of spatiotemporal scales. Therefore, accurate quantitative estimation of rainfall is an important input for engineering design applications where Precipitation Frequency Estimates (PFE) are highly sought [1]. The purpose of a precipitation frequency analysis is to determine the frequency at which certain intensities or depths of precipitation are expected to occur. Probabilistic modeling and statistical analysis techniques of extreme rainfall are used to provide PFE information and characterize the relationships between three important precipitation variables: intensity (or depth), duration, and frequency [2]. Such relationships are usually referred to as Intensity-Duration-Frequency (IDF) or Depth-Duration-Frequency (DDF) curves. Statistics derived from IDF or DDF curves are typically used

to develop design storms, which are then used as an input for a variety of engineering applications such as design of dams, levees, reservoirs, and urban sewer systems [3].

Precipitation frequencies are typically estimated using sparse gauge observations. The evolution of weather radars allows the spatially continuous estimation of rainfall at small temporal sampling intervals, thereby filling the observational gap of rain gauges in space and time. Radar does not measure surface rainfall directly; instead, it measures the backscattered power from the hydrometeors aloft and the received power is then converted into rainfall estimates with inherent errors. The availability of NEXRAD Quantitative Precipitation Estimates (QPE) in high temporal and spatial resolutions covering the United States (US) has motivated researchers to study the applicability of the radar-based QPE in deriving precipitation frequencies [4–8]. For instance, Overeem, et al. [8] used radar data covering the entire land surface of the Netherlands for a 10-year period (1998–2008) to derive radar-based areal reduction factors (ARFs), which were found comparable to those based on high-density rain gauge networks and thus concluded that radar data, after careful quality control, are suitable to estimate extreme areal rainfall depths.

For sites that sufficiently have long records with respect to the return period of the extreme precipitation quantile of interest, at-site frequency analysis can be an adequate approach. However, for un-gauged sites, or for sites with historical records that are too short to make a reliable prediction of extreme quantiles, data augmentation from neighboring sites is needed. Thus, two main approaches for the frequency analysis have been discussed in the literature. The first is an at-site estimation approach, which simply uses data at each station, while the second method is a regional estimation approach that makes use of observations from gauges sharing a homogenous region with similar climatological and physical characteristics [9–14]. Svensson & Jones [13] reviewed the different estimation methods of rainfall frequency analysis in nine countries and reported that, while each country's method is different, most of them use some form of regionalization to transfer information from surrounding sites to the target location. A regionalization method combines a local estimate of an index variable (typically the mean or median annual maximum rainfall) with a regionally derived growth curve to obtain a design rainfall estimate. Naghavi & Yu [15] applied a regional frequency approach to precipitation data in Louisiana using Annual Maximum Series (AMS) extracted from 25 synthesized stations with long periods of record. The results showed that the regional approach can substantially reduce the relative root-mean-square error (RRMSE) and the relative bias (RBIAS) in precipitation quantile prediction.

Although radar QPE can provide site (pixel)-specific PFEs with a high spatial resolution, regionalization techniques could be advantageous to reduce the sampling variability in the radar PFEs [4,6,16]. Eldardiry et al. [16] tested three regional estimation procedures and indicated lower uncertainty bounds associated with regional approaches compared to pixel-based PFEs. However, they also reported on the effect of the relatively short radar records on the uncertainty associated with the radar-based quantiles. Using QPE data during 1993–2000 over the Arkansas-Red Basin River Forecast Center (ABRFC), Durrans et al. [4] concluded that data heterogeneities and the temporally-limited data records are major factors that hinder the development of depth-area relationships using radar-rainfall data.

In this study, we assess the robustness of a probability weighted regional spatial bootstrap approach to estimate precipitation frequencies using radar data. This method was proposed by Uboldi et al. [17] as a resampling approach for estimation of parameters of rainfall annual maximum series statistical distribution. Using the regional spatial bootstrap technique, we investigate two main issues that impact the use of radar-based QPE in deriving precipitation frequency estimates: (1) the typically short historical records of radar-based QPEs; and (2) the effect of outliers in precipitation maxima series that could possibly cause unrealistic spatial gradients in IDF relations. We assess the utility of the spatial bootstrap approach in alleviating such limitations and compare the PFEs from the regional bootstrap approach against estimates derived using an at-site (pixel-based) method and PFEs reported in a US

gauge-based Precipitation-Frequency Atlas [18]. The study is performed using radar-based QPEs over the state of Louisiana, USA.

2. Datasets and Methods

2.1. Radar MPE Dataset

NEXRAD (Next-Generation Radar) is a network of more than 160 high-resolution S-band Doppler weather radars operated by the US National Weather Service (NWS). The NEXRAD system provides high-quality, high-resolution precipitation estimates for a wide range of hydro-meteorological applications [19]. Starting in 2002, the NWS implemented a processing algorithm called the Multisensor Precipitation Estimator (MPE) [1], which is currently used at the NWS River Forecast Centers (RFC) to produce a set of regional rainfall products using single- or multi-sensor analysis techniques. However, radar-based estimates can be highly uncertain due to a number of sampling and algorithm factors e.g., [20–22]. Therefore, the MPE algorithm applies bias-correction and co-Kriging optimal merging techniques using gauge reports [1,23].

The specific radar product used in this study is the operational product produced at the NWS Lower Mississippi River Forecast Center (LMRFC). This product is developed using the MPE algorithm, but also benefits from manual adjustments by human forecasters [22,24,25]. The MPE product has an hourly resolution and projected spatially using the Hydrologic Rainfall Analysis Project (HRAP) grid [26], with an approximate pixel size of 4-km x 4-km. The domain of the current study covers the entire state of Louisiana in the south-central US. With proximity to the Gulf of Mexico, different synoptic weather patterns are responsible for extreme events over Louisiana, including tropical storms, fronts, and convective air mass thunderstorms [27]. For the purposes of this study, the radar pixels covering Louisiana were extracted from the full LMRFC MPE product, resulting in a matrix with a total number of pixels of (180×140). The dataset comprises a total of 11 years covering the period of 2002–2012. To perform a precipitation frequency analysis, the annual maximum series (AMS) were extracted over each of the (180×140) radar pixels, resulting in a sample composed of a spatial field of 11 maxima representing the maximum hourly rainfall during 2002–2012 at each 4-km $\times$ 4-km pixel within the study domain.

2.2. Estimation of Parameters of AMS Probability Distribution

Various distributions have been proposed for modeling extreme events, including the Generalized Extreme Value distribution (GEV), Generalized Pareto distribution (GP), gamma distribution, lognormal distribution, among others. The GEV distribution was recommended for flood frequency analysis in the United Kingdom (UK) Flood Studies Report [28]. According to the gauge-based Precipitation-Frequency Atlas of the US National Oceanic and Atmospheric Administration (NOAA) Atlas 14 [18], the GEV distribution provided an acceptable fit to data more frequently than any other distribution and was chosen to model the annual maximum series of all the stations covering the US southeastern states (Alabama, Arkansas, Florida, Georgia, Louisiana and Mississippi). These conclusions were obtained using a goodness-of-fit test based on L-moment statistics for 3-parameter distributions along with the results of χ^2 and Kolmogorov-Smirnov tests and visual inspection of probability plots. Naghavi & Yu [15] examined six different distributions for extreme precipitation over Louisiana and concluded that GEV distribution outperforms other distributions. Therefore, the GEV will be adopted in the current study to represent the AMS. The GEV distribution is a three-parameter distribution developed within the extreme value theory and combines three different models: Gumbel, Frechet and Weibull distributions, which are often referred to as Types (I), (II), and (III) distributions, respectively. The probability density function of the GEV distribution, in terms of its three parameters: Location (α), Scale (β), and Shape ($\kappa \neq 0$), can be formulated as follows:

$$f_X(x) \;=\; \frac{1}{\beta}\left[1 - \frac{\kappa}{\beta}(x - \alpha)\right]^{(1/\kappa - 1)} F_X(x) \tag{1}$$

In this study, the method of linear moments is used for the estimation of the GEV distribution parameters. The method of L-moments offers several advantages over other methods (e.g., method of moments, and method of maximum likelihood), especially in the cases of small sample sizes [29,30]. The L-moment estimators for the GEV distribution parameters are given as follows:

$$\hat{\kappa} = 7.8590c + 2.9554c^2 \tag{2}$$

$$c = \frac{2}{(3 + \hat{\tau}_3)} - \frac{\log(2)}{\log(3)} \tag{3}$$

$$\hat{\alpha} = \hat{\lambda}_1 - \frac{\hat{\beta}}{\hat{\kappa}}[1 - \Gamma(1 + \hat{\kappa})] \tag{4}$$

$$\hat{\beta} = \frac{\hat{\lambda}_2 \hat{\kappa}}{(1 - 2^{-\hat{\kappa}})\Gamma(1 + \hat{\kappa})} \tag{5}$$

where τ_3 (or L-skewness) is the ratio of the third and second L-moment (a measure of skewness), λ_1 is the first L-moment (measure of distribution mean), and λ_2 is the second L-moment (measure of the scale or dispersion). Accordingly, the quantiles corresponding to different return periods, T, (e.g., T = 5, 10, 50, 100 years) can be estimated as follows:

$$x(q) = \alpha + \frac{\hat{\beta}}{\hat{\kappa}}\left[1 - \left(-\ln q\right)^{\hat{\kappa}}\right] \tag{6}$$

where q is the cumulative probability of interest that can be related to the return period T as $\left(q = 1 - \frac{1}{T}\right)$.

2.3. At-Site and Regional PFE Estimation Methods

Two estimation approaches for the frequency analysis of extreme precipitation, at-site (or pixel-based) and regional, are applied using the radar product and compared to each other. Uncertainties due to sampling effects are quantified in terms of confidence bounds using the difference between the 95th and 5th percentiles. The results of each method will also be compared to the corresponding gauge-based estimates that are reported in the NOAA Atlas 14 PFEs [18].

2.3.1. Pixel-Based Method

This approach is analogous to at-site frequency analysis of extreme precipitation from rain gauge stations, which was originally applied by the NWS to establish the rainfall frequency isohyetal maps for the US [31]. Treating each 4-km × 4-km HRAP radar pixel as a single station, this is equivalent to considering the domain of study as a dense network of stations that are located 4 km apart from each other. At each pixel, the AMS sample of hourly precipitation, constructed from the 11-year radar dataset, is fitted to the GEV model and the parameters of the distribution are estimated at each pixel using the L-moment method. The quantiles corresponding to different return periods (e.g., 5, 10, 25, 50, 100) can then be estimated at each pixel using the GEV parameters. Confidence intervals for the pixel-based parameter and quantile estimates are constructed using the classical scalar bootstrap procedure suggested by Efron [32]. The bootstrap procedure is used to generate a large number of samples (500 in our case) for each individual pixel.

2.3.2. Regional Spatial Bootstrap Method

This is a probability weighted regional method that was originally proposed by Uboldi et al. [17] as a resampling approach for estimation of parameters of the AMS distribution. This technique is based on the generation of a regional sample at any desired location by taking into account data observed at surrounding stations but with decreasing importance when distance increases. Thus, the probability of contribution of a certain station decreases as it goes far away from the desired location. The probability of sampling also takes into consideration the length of the time series at each station, and as such,

the possibility of oversampling can be avoided, and the use of short time series is enabled. This method is basically a spatial bootstrap technique in which a regional sample is generated repeatedly from the surrounding locations (pixels in the case of radar data) based on the randomness produced from the probability of data extracting. The procedure of this approach involves formation of a homogenous region, construction of a regional sample, estimation of statistical distribution parameters, repeating the regional sampling and parameter estimation several times as in any bootstrap technique, and finally obtaining a distribution of estimates for each parameter.

The regional sample of size (N) is constructed by extracting (N) observations randomly from all of the available data (M) in a homogenous region. The probability of extraction of each observation is assumed to be proportional to a prescribed Gaussian function (γ_m) of the distance between the station at the desired location (X) and any other station (K_m). Using spatially continuous radar observations (pixel resolution of ~4 km in the current study), the spatial bootstrap methodology is implemented as follows. For each pixel at a desired location X, and by prescribing distance-dependent extraction probabilities, observations from nearby pixels are selected more often than observations from pixels located far away. The probability of extraction of the mth observation located at a pixel (K_m) is given by the following relation:

$$\gamma_m = \exp\left\{-\frac{1}{2}\left[\frac{d_h(X, k_m)}{D_h}\right]^2\right\} \cdot \exp\left\{-\frac{1}{2}\left[\frac{d_v(X, k_m)}{D_v}\right]^2\right\} \tag{7}$$

where $d_h(X, K_m)$ and $d_v(X, K_m)$ are the horizontal and vertical distance between pixel K_m and the pixel at the desired location (X). The D_h and D_v are scale parameters that are selected to impose some degree of smoothing and were chosen in this study to be equal to the standard deviation of the available distances between (X) and (K_m).

Normalized by the sum of probabilities of all the observations (M), the probability of extraction of each observation from N set of available observations can be obtained as follows:

$$\overline{\gamma}_m = \frac{\gamma_m}{\Gamma}, \ \Gamma = \sum_{m=1}^{M} \gamma_m \tag{8}$$

By sorting the (M) observations in a descending order according to their probability of extraction ($\overline{\gamma}_m$) and assigning each observation a number (m) from 1 to M, a series of sequential ordered dataset is obtained. The cumulative normalized probability of extraction ($\overline{\gamma}_m$) of each observation ranges between (0, 1) and the probability of extraction of this cumulative probability is assumed to be uniformly distributed, i.e., ($\overline{\gamma}_m$)~U (0,1). A continuous random variable (ρ) is then used to implement a random number generator for a discrete random variable (m) with any prescribed (non-uniform) probability distribution on positive integers up to a generic M. By generating a random number (ρ), the corresponding cumulative probability ($\overline{\gamma}_m$) is equal to the generated random number (ρ) and realization number (m) is equal to the first observation that has cumulative probability greater than or equal to the generated probability (ρ).

The spatial bootstrap regional approach requires the formation of a homogenous region surrounding each pixel, from which a regional sample can be constructed. The identification of homogenous regions is a non-trivial step in the regional frequency analysis, and it may require subjective judgement [33]. A homogenous region is the area including a group of sites, or pixels as in the case of radar fields, that share similar physical characteristics. The advantages of working with a homogeneous region is that the historical data available within the region can be pooled to get an efficient estimate of parameters and hence a more robust quantile estimate [34]. Hosking & Wallis [33] strongly preferred to base the formation of homogenous regions on site characteristics (e.g., by using geographical delineation, cluster analysis, or principle components analysis) and to use the at-site statistics only in subsequent testing of the homogeneity of the proposed set of regions. Conventional regionalization techniques identify a fixed set of sites to form a contiguous region, resulting in fixed-boundary regions without smooth transitions.

Burn [10] presented the Region of Influence (ROI) approach for defining homogenous region, in which every site can have a potentially unique set of gauging stations to be used in the estimation of at-site extremes. The ROI technique is recommended as it avoids the transition problems across fixed boundaries by introducing smooth change in the estimates across the boundaries of the regions. The selection of the radius of influence is a trade-off problem, in which a large radius R will increase the number of sites included in each ROI, but at the expense of the homogeneity of the set of sites included. Conversely, a small radius R will ensure the homogeneity of the sites included, but the information transfer will be decreased due to the smaller number of sites. In this study, the ROI approach is applied by using a square window with an area equivalent to $(2R + 1)^2$ bounding the pixel of interest (R is the radius of influence and is used here to refer to the number of pixels considered in the horizontal or vertical directions between the central pixel and the edge of the square window). The window forms a homogenous region and constructs the regional sample for the target pixel (central pixel) using the pixels lying inside this window.

Since the choice of a homogenous region, or the window size, should be based on climatic and physical characteristics, the US Climate Divisions are used in this study to provide an indication for the reasonable range of the radius of influence (R). Louisiana has nine Climate Divisions (Figure 1) and the average area of each climate division is approximately covering a window with side length of about 31 pixels, which corresponds to R = 15 pixels.
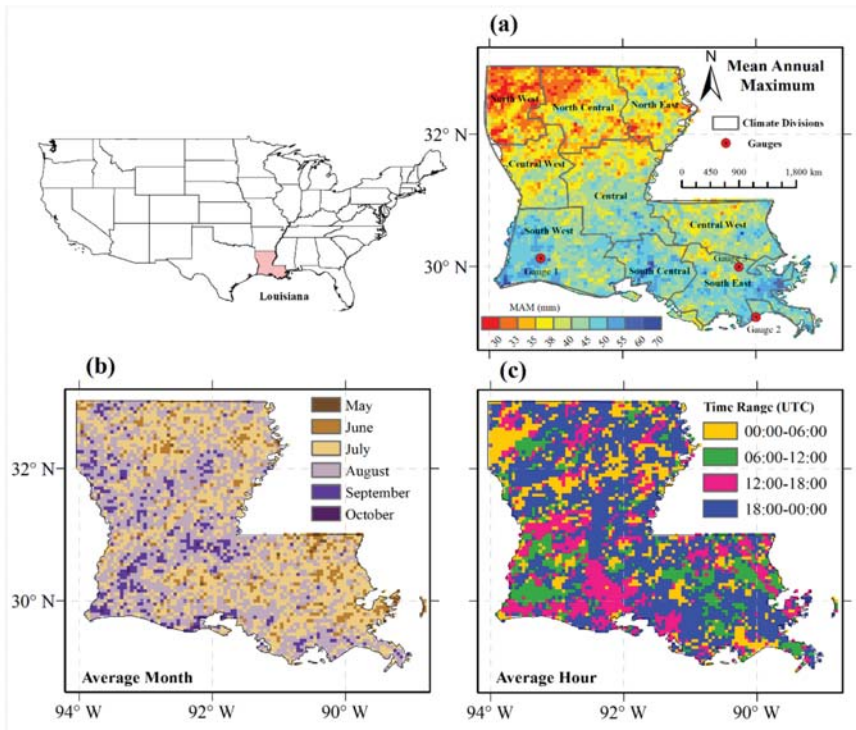
Figure 1. (**a**) Spatial distribution of the hourly Mean Annual Maxima (MAM; in mm). (**b**) The average month of AMS occurrence in each pixel. (**c**) The average 6-h of AMS occurrence in each pixel. Results are based on the period of study (2002–2012).

Therefore, we chose R = 15 as a threshold for identifying the homogenous regions to estimate PFE. In this study, we tested different square windows ranging from R = 3 pixels to R = 15 pixels (results are only shown for R = 5 and 10) to study the effect of the region size on the uncertainty of the estimates. For example, increasing the window size to 21 × 21, by setting R = 10 pixels, allows for many more pixels (M = 441 × 11; 441 pixels with AMS of 11 observations in each pixel) to be included in the region of each target pixel. The scale parameters in Equation (7), i.e., D_h and D_v, are chosen to be approximately equal to the standard deviation of each radius of influence (for R = 5, $D_h = D_v = 1$ pixel and for R = 10, $D_h = D_v = 3$ pixels). The regional sample size is chosen to be the same as the actual number of years available in the radar dataset, i.e., N = 11 (sampled out of the M observations).

In order to reduce the likelihood of extracting annual maxima that might come from the same event, a constraint is added in such a way that the gap in the time stamp of any two annual maxima extracted from two different pixels must be greater than 6 h. This criterion is evaluated using the Julian Date in which the 6 h represent 0.25 day. For instance, if the extracted annual maximum occurs in a certain Julian Date (JD), then any new annual maximum must have a new Julian Date greater than (JD + 0.25) or smaller than (JD − 0.25). This restriction might not be necessary in the case of gauge-based PFE analysis since the gauges are separated with relatively large distances, and therefore it is less probable to have annual series in two gauges that share exactly the same events. On the other hand, the application of this conditioned extracted annual maxima is critical to the radar-based annual maxima, since they are provided on a uniform grid with high spatial resolution (4-km × 4-km in our dataset).

3. Results

3.1. Characterization of Annual Maxima

Radar precipitation estimates provide new possibilities to investigate the climatology of extreme rainfall at high spatial resolutions and over large areas [7]. Louisiana is considered one of the wettest of the contiguous 48 US states with extreme events that are generated by various rainfall mechanisms. Extreme events in the Southeastern US are typically generated from different synoptic weather patterns, for example, tropical storms, fronts, and convective airmass thunderstorms [27]. Figure 1a shows the Mean Annual Maxima (MAM) rainfall depth for each pixel in the domain of the study area. Most of the maxima are in the range between 20 mm and 100 mm with a significant spatial increase towards the gulf coastal zone. Figure 1b depicts the spatial distribution of average month of occurrence for annual maximum rainfall and shows dominance of the summer season (June—July—August) throughout most of the state. These results agree with Maddox et al. [35] who concluded that most of the extreme events that cause flash floods are of convective nature with the predominance of events in the warm season (April-September). The diurnal distribution of the annual maxima is illustrated by the average 6-h of annual maxima occurrence (Figure 1c). Most of the annual maxima occurred between 18:00 UTC and 00:00 UTC, while fewer number of events occurred in the two intervals (00:00 UTC–06:00 UTC) and (12:00 UTC–18:00 UTC).

3.2. Radar-Based PFE using Regional Sptail Bootstrap

Before presenting the results of PFEs using the spatial bootstrap technique, we first examine the traditional at-site method (pixel-based in case of using radar datasets). The pixel-based estimation procedure described in Section 2.3.1 was applied to the radar dataset to estimate the GEV distributional parameters and the corresponding PFE for different return periods ranging from 2 to 100 years. Confidence intervals for the estimated parameters and PFEs were also derived using classic scalar bootstrap sampling at each pixel. The parameters and PFEs for each pixel are represented using the mean of the 500 runs of bootstrap. The difference between the lower 5% and upper 95% quantiles of the bootstrap samples are used to quantify the uncertainty in the estimates. The confidence intervals are calculated using a non-parametric method, in which a probability is initially assigned to the sorted

values of the sample ((0.5/n), (1.5/n), ([n—0.5]/n)), where n is the sample size (n = 500 bootstrap runs). The quantiles are then computed by setting the probability to be equal to the confidence limit required, e.g., 0.95, 0.90, 0.05, or 0.1. The first and last value in the bootstrap sample are assigned to the quantiles for probabilities less than (0.5/n) and greater than ([n—0.5]/n), respectively. Figures 2 and 3 show the GEV parameters estimated at each pixel using the pixel-based and regional spatial bootstrap methods, respectively, over the domain of study covering Louisiana.
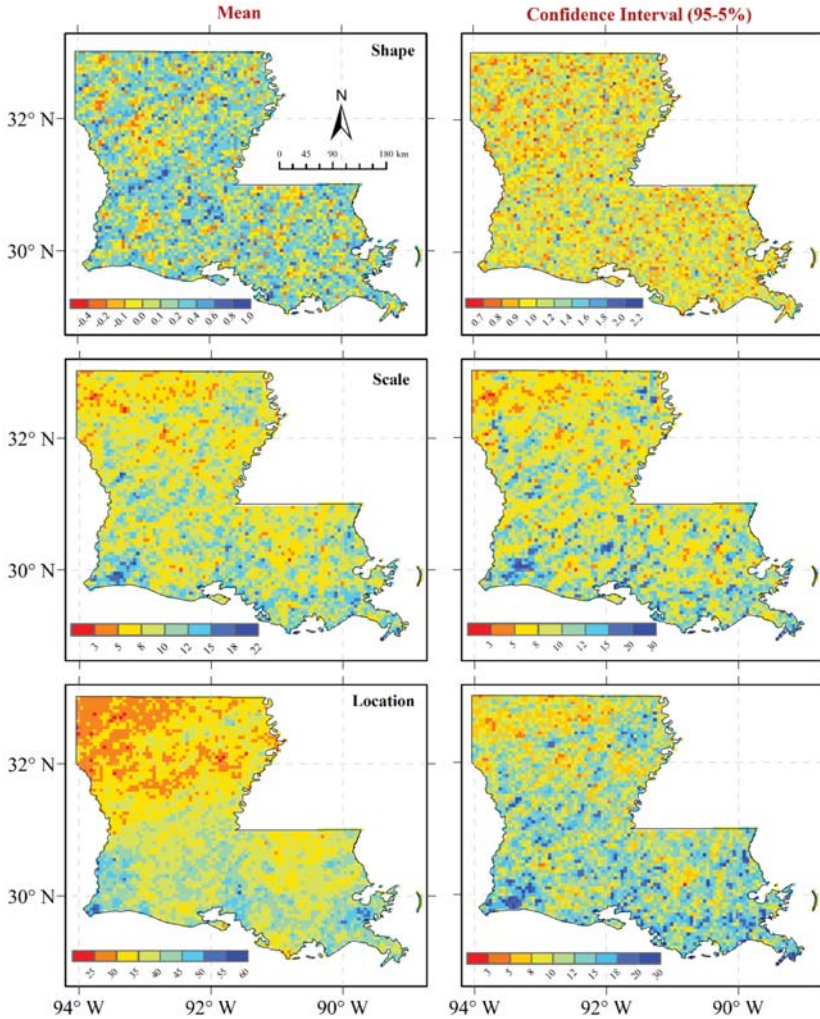
Figure 2. The GEV distribution parameters (shape, scale, and location parameters) from the pixel-based approach. Left Panels: Mean of 500 bootstrap runs. Right Panels: The confidence width (95–5% percentiles).
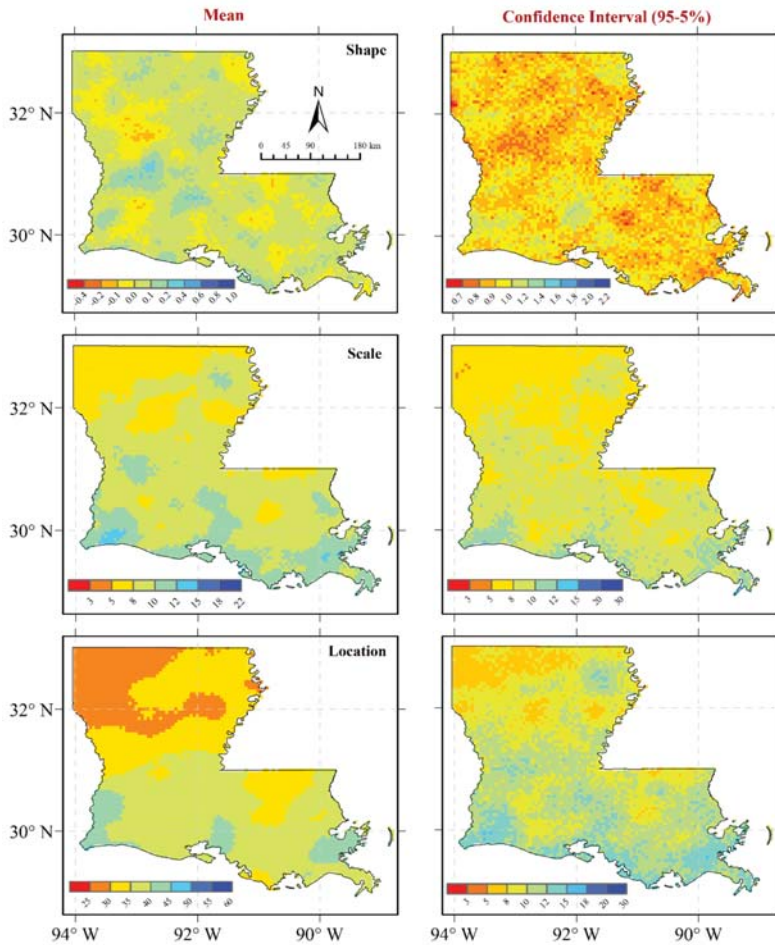
Figure 3. The GEV distribution parameters (shape, scale, and location parameters) from the spatial bootstrap (region-based) approach. Left Panels: Mean of 500 bootstrap runs. Right Panels: the confidence width (95–5% percentiles).

The shape parameter, estimated from the average of 500 bootstrap runs, varies between positive and negative values mostly between [−0.5, 0.5]. The 5% and 95% confidence of the shape parameter have values below −0.5 and above 1 due to the sampling variability. The scale parameter, in most pixels, falls in the range between 5 and 20, with some subtle spatial patterns. The location parameter has noticeable spatial gradients similar to those of the MAM (Figure 1a) where the location parameter increases from north to the south and as we get closer to the Gulf boundary. The sampling effect on both of the scale and location parameters is evident in the 5% and 95% confidence limits. The corresponding PFEs are displayed for two representative return periods of 2 and 10 years (Figures 4 and 5). The PFE results show significant variability in space with clear gradients from north to south. The uncertainty associated with these estimates is fairly large, especially for large return periods, e.g., 50 and 100 years (figures not shown). The spatial maps also show clear signs of irregularities and inconsistency in the spatial variability of the estimated quantiles, which are mostly noticed for large return periods.
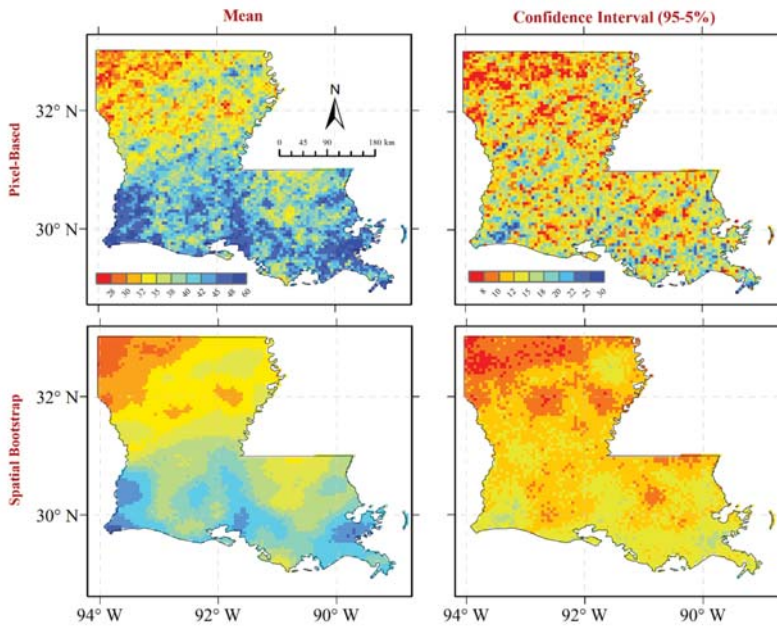
Figure 4. The rainfall depth (in mm) and the confidence width (95–5% percentiles) corresponding to 2-year return period from the pixel-based (upper panels) and spatial bootstrap approaches (region-based) (lower panels).
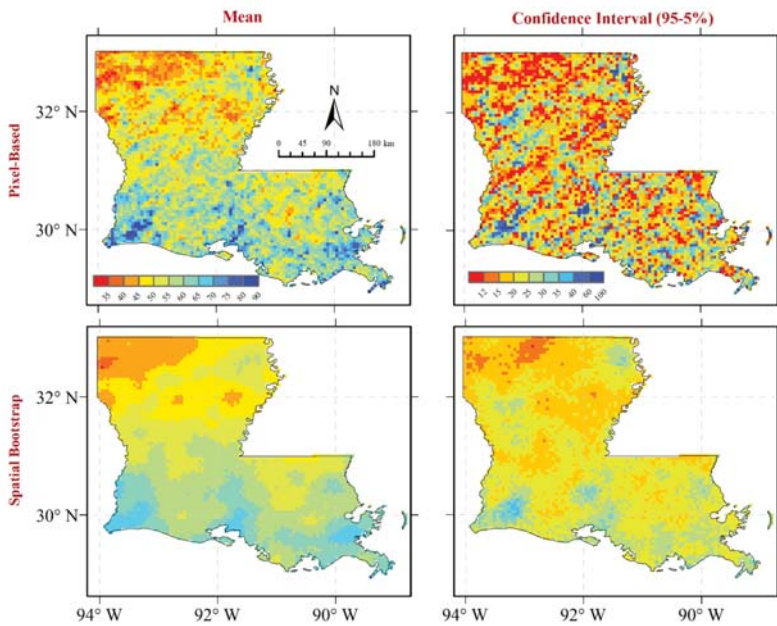
Figure 5. The rainfall depth (in mm) and the confidence width (95–5% percentiles) corresponding to 10-year return period from the pixel-based (upper panels) and spatial bootstrap approaches (region-based) (lower panels).

The confidence limits are estimated using the spatial bootstrap technique for 500 runs using a moving window of 11 × 11 pixels (R = 5). Compared to the pixel-based approach, the results suggest that the spatial bootstrap approach reduced the estimated parameters and resulted in narrower confidence intervals. For instance, the mean shape parameters, in most of the pixels, went down to the range [−0.2, 0.2] with a noticeable reduction in the width of the uncertainty bounds. The reduction in the dispersion of the estimated parameters is attributed to the gain from the repeated sampling from the surrounding pixels, which is the main advantage of a regional estimation as opposed to using information available at each pixel only. Sampling from a homogenous region resulted in smoother fields of the GEV parameters with less sampling variability. Because of the short record available in each pixel, only 11 years, the pixel-based estimation varies considerably from one pixel to another, which was circumvented when using the regional spatial bootstrap estimation with the moving window at each pixel. Increasing the size of the moving window to (21 × 21) or R = 10 pixels resulted in lower variability and more smoothness for the estimates transition between the pixels (figure not shown), but possibly at the expense of losing details in the spatial patterns.

Figures 4 and 5 display the PFEs using the GEV distribution parameters for return periods of 2 and 10 years. Improvements in the smoothness of the different PFEs can obviously be seen when using the spatial bootstrap approach over the pixel-based approach. The smoothness in the PFEs patterns by the spatial bootstrap resembles to a great extent the smoothing algorithm performed by Durrans et al. [4] who used simple distance-weighted averaging procedures to spatially smooth the estimates of sample L-moments. Their smoothing algorithm reduced the effects of sampling variations caused by the short time series used, only eight years in their study.

3.3. Comparison Against Gauge-Based PFE

In this section, the NOAA Atlas 14 gauge-based PFEs [18] are contrasted against the corresponding frequencies estimated using the two approaches presented earlier; pixel-based and spatial bootstrap regional estimation methods. The gauge-based AMS used in the Atlas 14, as well the corresponding PFEs with their 90% confidence intervals, were acquired from the NOAA's Hydrometeorological Design Studies Center (HDSC) web-based data server. We used the gauge-based PFE from the NOAA Atlas as a reference to assess the robustness of the spatial bootstrap method when (a) estimating PFEs with short radar samples, or (b) in cases of having outliers in the radar AMS sample. However, it is important to note that this comparison does not imply that PFEs from gauges are the true estimates, simply because they also have their own uncertainties caused by sampling variability and the estimation process itself [36]. Nevertheless, the comparison will provide some insights into the performance of the regional spatial bootstrap method in deriving PFEs using radar-based estimates.

The NOAA Atlas 14 applied a regional frequency analysis approach that is different from the spatial bootstrap technique used in the current study. The main difference is in how the regional sample is constructed from the homogenous region formed for each station. In the Atlas 14 method, a homogeneous region is defined for each gauge by grouping the closest 10 stations. The 10 stations are then added to or removed from the region based on factors such as distance from a target station, elevation difference, difference in MAMs at various durations, and inspection of locations with respect to mountain ridges. The AMS for a network of 33 hourly gauges in Louisiana is retrieved from the HDSC and used in the current study to identify differences in the AMS constructed from the radar QPE versus those from the gauges (Figure 6).
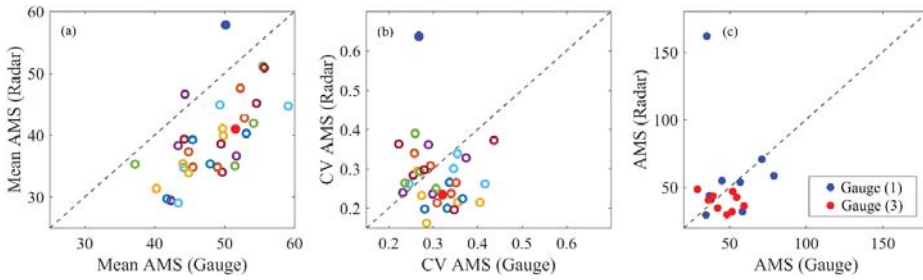
Figure 6. (**a**) the mean annual maxima rainfall depth extracted from NOAA Atlas 14 gauges and the corresponding radar-pixel (each color represents one of the 33 gauges in Louisiana retrieved from NOAA HDSC web-based data server). (**b**) same as (**a**) but reporting the coefficient of variation. (**c**) comparison of AMS from gauge data and radar-based estimates (for common period 2002–2010) at the location of two example NOAA Atlas-14 gauges (indicated in Figure 1a).

Figure 6a shows that the radar-based QPE product has an overall lower value (AMS) than the corresponding gauge-based AMS, with an average underestimation of 9 mm. Such underestimation of radar-based precipitation can be partially attributed to the areal estimation of precipitation in case of radar pixel as opposed to point gauge. However, given the high resolution of radar (4-km × 4-km), the effect of point-area discrepancies is negligible for small areas. For instance, according to the values given in TP-29 [37], the percent of area-to-point precipitation in case of hourly rainfall and for areas of less than 16 km^2 is higher than 95%. In terms of the variability of the AMS, 20 gauges experience higher coefficient of variation (average = 0.34) compared to the corresponding radar pixels (average = 0.25). A higher variability in the gauge-based AMS is attributed to longer record available (with an average of 38 years for gauges in Louisiana) compared to only 11-years of radar-based AMS that are used in this study.

Three representative gauges (from the NOAA Atlas 14), are selected for further comparison analysis (Table 1).

Table 1. The NOAA Atlas 14 gauges selected in our study to evaluate the spatial bootstrap approach when using radar-based precipitation estimates.

Gauge	Latitude	Longitude	NOAA Atlas14 AMS Size
Gauge (1)	30.12°	−93.23°	49 years
Gauge (2)	29.23°	−90.00°	26 years
Gauge (3)	29.99°	−90.25°	64 years

The three gauges are located in the southwest and southeast climate divisions of Louisiana (Figure 1a). Figure 6c shows plots of the AMS extracted from gauge (1) and the coincident pixel for 9 years covering the common period (2002–2010). The gauge-based AMS is available for 49 years from 1962 to 2010 which is a long record compared with the 11-year radar QPE data used in the current study (2002–2012). It is noted that the 2003 annual maximum from the radar QPE is much higher compared to that of the corresponding gauge, which suggests that this particular value might be an outlier. The mean and standard deviation of the AMS for this pixel are 58 mm and 37 mm respectively. When excluding the outlier observation, the standard deviation for the AMS is only 14 mm, which indicates the high variability that might result from this individual value. Moreover, upon applying the Grubbs-Beck (GB) outlier detection test, this observation is considered an outlier at a 5% level of significance.

Figure 7 shows the effect of including annual maxima events from neighboring cells by comparing the range of AMS (minimum and maximum hourly precipitation in AMS sample) at the location of

three gauges in Louisiana (indicated in Figure 1). Again, the outlier identified at the pixel coincides gauge (1) is reflected in the range of AMS formed when using regional approach. In addition, it is very evident in gauges (2) and (3) that sampling from neighboring cells, i.e., forming a regional sample, can significantly increase the range of extreme precipitation events that are not captured in the 11-year AMS sample of the pixel. For example, in case of gauge (3), the AMS sample from a region of R = 10 can range between 16.8 mm and 109.5 mm, which covers the range of gauge-based AMS (between 28.4 and 102 mm). When using a pixel-based approach at the same gauge location, sampling from an 11-year AMS can only ranges between 29.9 mm and 63.1 mm.
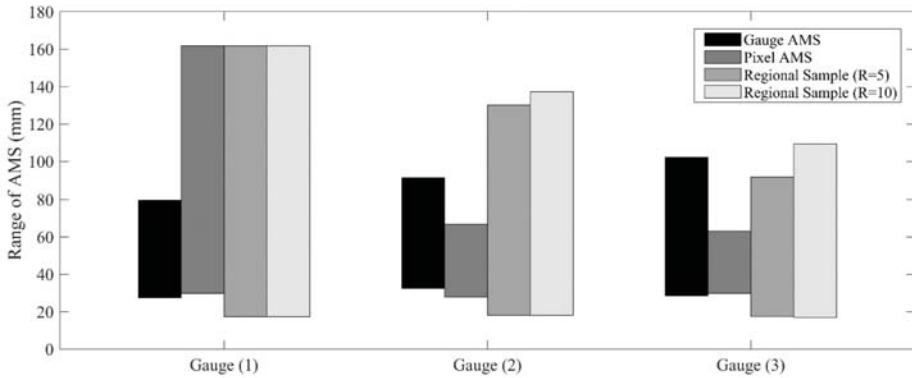
Figure 7. The range of AMS from gauge data, radar pixel, and radar-based regional sample considering a radius of 5 and 10 pixels. Each bar ranges between the minimum and maximum value in AMS sample extracted at the location of gauge (see Figure 1a for gauges locations).

To assess the impact of outliers, we estimated the PFE quantiles at the specific location of gauge (1) using both the pixel-based and the spatial bootstrap methods. The unusually high radar AMS value (in 2003) resulted in a rather higher mean PFEs and wider uncertainty bound when using the pixel-based approach (Figure 8), while the spatial bootstrap technique was much less influenced by it.
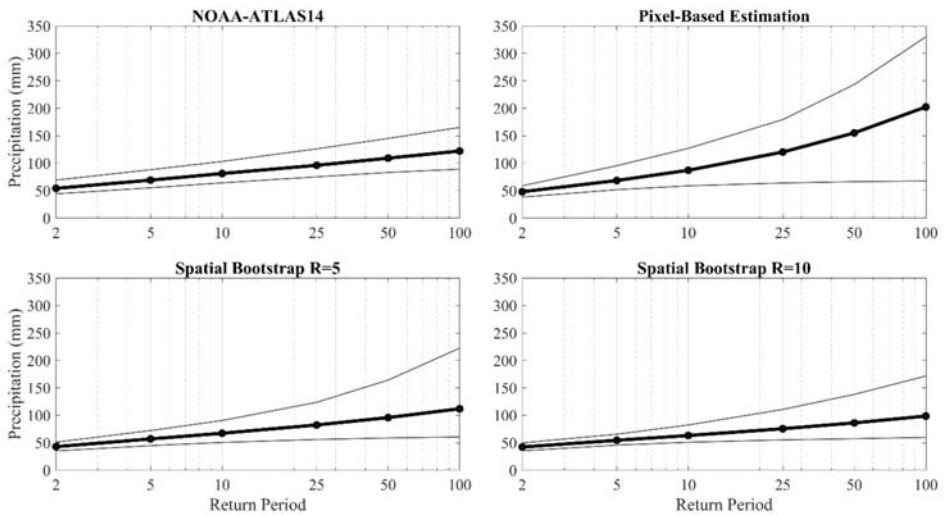
Figure 8. Precipitation Frequency Estimates (PFE) and (95–5%) confidence limits using different estimation approaches at the location of Gauge (1).

For example, for 25-year return period, the pixel-based estimation resulted in a mean PFE of 120 mm compared to 96 mm and 82 mm estimated in Atlas 14 (using gauges) and spatial bootstrap (using AMS from radar pixels with R = 5 pixels), respectively. The (95–5%) percentile difference of 25-year return period dropped down from 115.5 mm when using the pixel-based estimation to 68 mm using spatial bootstrap with a radius of influence R = 5 pixels (compared to 51 mm from the regional approach adopted in Atlas 14). Further reduction in mean PFEs and confidence limits are also noticed in Figure 8 when using spatial bootstrap while augmenting the regional sample using a larger radius of influence (e.g., R = 10). The reduced effect of possible outliers is one of the benefits of the spatial bootstrap technique since the combined use of multiple pixels enables reducing the impact of such very rare events (Uboldi, et al., 2014). Owing to longer records available for precipitation in most of the gauges (49 years for Gauge 1), the gauge-based PFEs from Atlas 14 have narrower uncertainty bounds for larger return periods (>10-year). Overall, the PFE results using the spatial bootstrap method are closer to those of the Atlas 14 than the pixel-based approach. For example, the spatial bootstrap method (with R = 5 pixels) resulted in an overall mean absolute percentage deviation (assuming Atlas 14 as our reference) of 15% compared to 25% when using pixel-based approach. When compared to NOAA Atlas 14 PFEs, the spatial bootstrap method resulted in smaller confidence intervals for shorter return periods (less than 10 year-return period).

Gauge (2) represents an example where AMS sample is constructed for a different period (1948–1981) compared to radar (2002–2012). Unlike gauge (1), the precipitation frequencies estimated by the NOAA Atlas 14 approach are quite larger than those estimated by the radar QPE dataset when using the pixel-based estimation. Figure 8 shows lower mean estimates for the quantiles and very narrow confidence intervals in the pixel-based estimation compared to Atlas 14 regional estimation method. For example, 5-year PFE using pixel-based estimation resulted in a confidence interval of 16 mm compared to 36 mm from gauged-based PFE. The lower quantiles estimates can be attributed to an overall underestimation in radar-based AMS values as compared to those estimated by gauge (average underestimation in mean annual maxima is 14.4 mm), while the less variability is due to the small standard deviation of the AMS. This narrow confidence bounds discloses one of the limitations of using the conventional bootstrap resampling with small sample sizes, since it will never generate an observation either larger or smaller than the maximum or minimum AMS observation [38].

When applying the spatial bootstrap technique over Gauge (2), it resulted in lower PFEs estimates compared to the NOAA Atlas 14 estimates. While the mean PFEs from both the pixel-based estimation and the regional spatial bootstrap techniques are very comparable, the spatial bootstrap resulted in wider confidence intervals. This is attributed to the addition of observations, other than those included in the pixel sample (Figure 7), that introduced more variability to the quantile estimates in a way that makes them closer to those derived by the gauge-based PFE from NOAA Atlas 14. Unlike the expected reduced variability in regional-based approach, the variability increased with opening the moving window to larger size (Figure 9; using R = 10) to take advantage from more pixels surrounding the pixel that coincides with Gauge (2). For example, 50-year return period PFE has a higher (95–5%) confidence width of 67 mm when using the spatial bootstrap (R = 10 pixels) compared to only 24.4 mm from pixel-based estimation. An underestimation in radar-based PFEs can be attributed to the conditional bias that is typically manifested in radar QPE products [16]. The conditional bias characterizes the performance of QPE products at different ranges of the rainfall amount [39]. An increase in the QPE bias at high rainfall rates [40] propagates in the radar-based PFE analysis regardless of the PFE estimation method and results in an overall underestimation of the PFE quantiles as illustrated in Figure 9c.
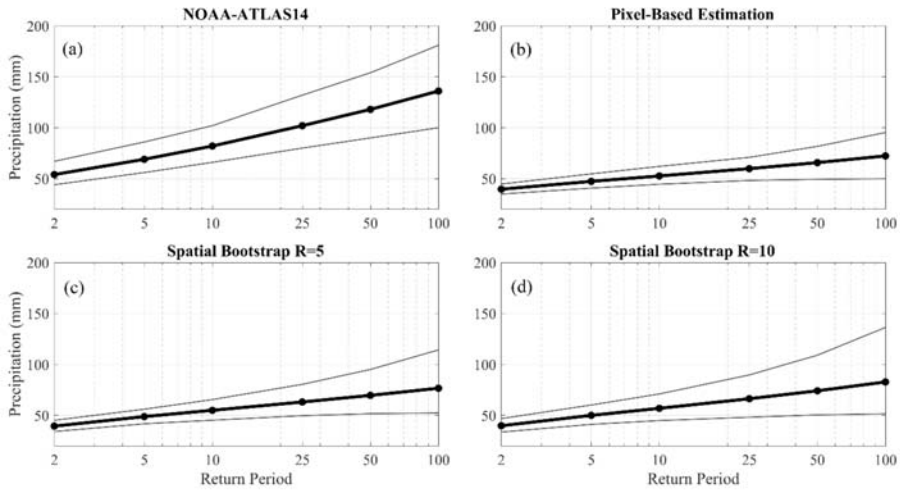
Figure 9. Precipitation Frequency Estimates (PFE) and (95–5%) confidence limits at the location of Gauge (2) based on (**a**) NOAA-Atlas14, (**b**) pixel-based approach, and (**c**,**d**) regional spatial bootstrap with a region of R = 5 pixels (**c**) and R = 10 pixels (**d**).

3.4. Effect of Regional Sample Size

In our analysis, we opted to use the same sample size of the radar record (11 years) to highlight the differences in the estimation methods, i.e., pixel-based estimation vs. The spatial bootstrap approach. However, since one of the advantages of the spatial bootstrap technique is the ability to incorporate more information from neighboring pixels, it is important to assess the effect of the size of the regional sample. Figure 10 compares the percentage change in the mean quantiles and the confidence interval (95–5%) of PFEs when using pixel-based approach and regional spatial bootstrap methods as compared to our reference (i.e., NOAA Atlas 14).
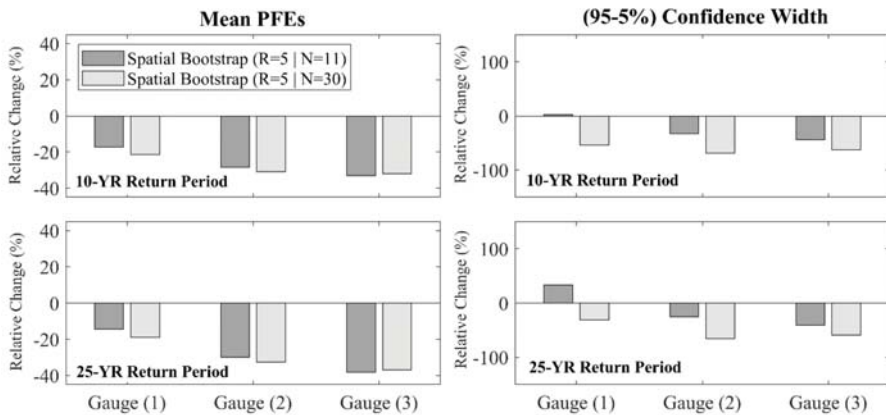
Figure 10. Percentage change in mean quantiles and confidence interval (95–5%) when using the spatial bootstrap method with sample size of 11 and 30, compared to NOAA Atlas 14 gauge-based PFES (only 10 and 25-year return periods are shown).

The results are presented at the locations of the three selected gauges in southern Louisiana (indicated in Figure 1). The spatial bootstrap method produced lower mean quantiles

(i.e., negative changes) for the three gauges with slight differences when increasing the regional sample size. For example, at the location of gauge (2), a regional sample of size N = 11 would underestimate (as compared to NOAA Atlas 14) the mean 10-year PFE by 28.4% (compared to 30.8% in case of N = 30). Increasing the regional sample of the spatial bootstrap can improve the PFEs estimation in case of gauge (3). The relative change in the mean 25-year quantile for the pixel coincides gauge (3) slightly reduced from −38% (N = 11) to −37% (N = 30). In terms of uncertainty, increasing the regional sample size can result in narrower confidence intervals as compared to gauged-based PFEs. For example, when using a regional sample size of N = 30 at Gauge (3) location, 25-year PFE has a narrower confidence interval with a reduction of −58.9% compared to the gauge-based PFEs (as opposed to a −40.8% reduction when using a sample size of N = 11).

4. Discussion

Accurate information on PFEs are critically needed at high temporal and spatial resolutions to serve in various water resources planning and design purposes. The estimation of PFEs becomes challenging when dealing with short data records to derive precipitation frequencies for large return periods [30]. Using a short sample size to fit the extreme value distributions results in large uncertainties when estimating distribution parameters and quantiles, especially for short durations, e.g., hourly PFEs. Therefore, implementing a regional frequency analysis is an effective means for trading space with time. However, when using precipitation estimates from remote sensing data, e.g., radar or satellite products, applying a robust regional frequency analysis is driven by: (1) accurate estimation of extreme values; and (2) definition of a homogenous region. This study investigated the utility of the spatial bootstrap technique as a potential regional approach to derive precipitation frequencies using radar-based precipitation datasets that typically have short observational records. The spatial bootstrap approach has the advantage over pixel-based estimation to augment the sample size by sampling from a homogenous region surrounding the pixel of interest. Our results indicated that the spatial bootstrap technique can provide spatially smoother distribution parameters and associated quantiles compared to the pixel-based approach, which reduces the unrealistically high variations between neighboring pixels over the fine-resolution radar grid (4-km × 4-km in the case of Stage IV).

Defining the spatial extent of a homogenous region is an important factor to consider when using a spatial bootstrap technique. The selection of the region size is a trade-off problem, in which larger regions will increase the number of pixels and the overall sample size, but at the expense of the homogeneity of the pixels included in the analysis. A larger region will also result in a reduction of the uncertainty of the PFEs (Figure 8). As recommended by [33], it is strongly preferred to base the formation of homogenous regions on site characteristics, using for example geographical delineation, cluster analysis, and principle components analysis. For example, a square region of pixels, as used in our study, might not be appropriate in case of complex terrain. Therefore, in such cases, a careful selection of a homogenous region should include different attributes of the study region such as physiographic catchment characteristics, geographical location attributes, and meteorological factors [41]. At-site or pixel-based statistics can be then used in subsequent testing of the homogeneity of the proposed set of regions.

Our tests on the effect of the regional sample size showed that a longer sample size can significantly reduce the uncertainty associated with large return periods, e.g., 10-year PFEs (Figure 10). When using radar data for PFE analysis, the regional sample size could be increased beyond the actual record length, but without significantly impacting the estimation of the mean PFEs. It is noted that the desired increase of the regional sample size might lead to over-sampling by including observations of similar events in the same synthetic sample; however, the spatial bootstrap method avoids such problems by assigning distance-dependent probabilities to individual observations, rather than to specific pixel sites. While our study focused on the frequency analysis of precipitation at hourly scale, the same regional approach can be implemented at longer durations, e.g., 6-h and 24-h, to derive DDF or IDF curves required for design purposes.

The results of the radar-based PFE were assessed versus those from the NOAA Atlas 14 that were developed using a gauge-based regional frequency analysis. The comparison indicated that pixel-based approach was highly sensitive to observational and sampling variability, and as such can yield much higher or lower PFE estimates compared to the gauge-based PFE. On the other hand, region-based spatial bootstrap approach was less sensitive to sampling effects and short records of radar data, thanks to its regional sampling mechanism. The spatial bootstrap technique provides more realistic representation of the PFE confidence intervals and thus can be considered more reliable when assessed against the reference NOAA Atlas 14 frequency estimates. Since spatial bootstrap technique is less sensitive to outliers, it can be more robust when applied using data that typically contain outliers in extreme precipitation, such as the case of most real-time radar products, including the Stage IV product [25]. The spatial bootstrap approaches are still prone to the systematic biases that are inherent to most radar-rainfall products. Conditional biases, which impact the extreme rainfall values and propagate into the PFE estimation process, need to be adjusted at the radar-rainfall estimation phase before being used for PFE applications. Isolating the effect of the inaccurate estimation of the extreme values by the radar product from other factors, e.g., selection of homogenous region or sample size, is beyond the scope of our study. A future work, e.g., through some simulation-based approach, can quantify how the systematic biases in extreme value estimation can mix with other factors and how they individually (and combined) affect the overall PFE results.

5. Conclusions

Traditionally, Precipitation Frequency Estimates (PFE) information is based on near-point observations of sparsely distributed rain gauges. The limited spatial availability of rain gauge stations, and their lack of areal representation, calls for exploring the utility of weather radar techniques for PFE analysis. This study examined the applicability of a spatial bootstrap regional approach to derive PFEs using radar-based Quantitative Precipitation Estimates (QPE). The focus was on whether the spatial bootstrap regional method can address typical limitations in using short-record radar datasets for PFE analysis. The analysis was performed over the domain of the state of Louisiana in southcentral USA. The key conclusions of our study are as follows:

1. The spatial bootstrap as a regional method can successfully alleviate the effect of short record availability in radar-based QPE (typically 10–20 years) by bootstrapping spatially from neighboring pixels to gain more information from a climatologically homogenous region.
2. The use of the spatial bootstrap regional method resulted in PFE quantiles and distribution parameter spatial fields that are smoother and less noisy compared to the pixel-based approach. Spatial gradients in the PFE quantiles are distinctly evident across the domain of the entire state.
3. Augmenting the sample size and/or the region of influence in the spatial bootstrap showed a significant reduction in the estimated uncertainty of the PFEs at different return periods.
4. Compared to a pixel-based approach, the spatial bootstrap technique is less sensitive to observational and sampling variability and can provide more realistic representation of the PFE confidence intervals. Thus, when compared with the gauge-based NOAA Atlas 14 frequency estimates, PFEs from spatial bootstrap method can be considered more reliable than pixel-based estimation. However, for some cases where QPE estimates have inherent systematic bias especially for extreme rainfall, both of the spatial bootstrap and pixel-based estimation methods resulted in considerable underestimation in PFEs.

The overall results of the current study indicate the potential power of regional spatial bootstrap technique in deriving PFEs from radar-based QPE at high spatial and temporal resolutions. Given the global coverage of satellite data at high spatiotemporal resolution, it is of interest, particularly in regions with scarce in-situ data, to advocate the use of satellite-based PFEs in the design, operation, and planning of infrastructure. Therefore, a robust regional approach, as the spatial bootstrap method, can be very useful in reducing uncertainties associated with satellite-based PFEs. Future studies can

also explore a viable approach that combines information from both radar and rain gauge sources to capitalize on their respective strengths and improve the PFE estimation process. Such accurate and regionally representative PFE information are critically needed for various water resources engineering planning and hydrologic design applications.

Author Contributions: H.E. conceived and designed the study, carried out data analysis and drafted the manuscript. E.H. participated in the design of the study, contributed to the interpretation of data and critically revised the draft of the manuscript providing substantial contributions to improve the presentation of results. Both authors read and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

References

1. Seo, D.-J.; Seed, A.; Delrieu, G. Radar and multisensor rainfall estimation for hydrologic applications. In *Rainfall: State of the Science*; Gebremichael, M., Ed.; American Geophysical Union: Washington, DC, USA, 2010.
2. Pecho, J.; Fasko, P.; Lapin, M.; Gaál, L. *Analysis of Rainfall Intensity-Duration Frequency Relationships in Slovakia (Estimation of Extreme Rainfall Return Periods)*; EGU General Assembly: Vienna, Austria, 2009.
3. Chow, V.T.; Maidment, D.R.; Mays, L.W. Applied Hydrology. In *Water Resources Handbook*; McGraw-Hill: New York, NY, USA, 1988.
4. Durrans, S.R.; Julian, L.T.; Yekta, M. Estimation of depth-area relationships using radar-rainfall data. *J. Hydrol. Eng.* **2002**, *7*, 356–367. [CrossRef]
5. Allen, R.J.; DeGaetano, A.T. Considerations for the use of radar-derived precipitation estimates in determining return intervals for extreme areal precipitation amounts. *J. Hydrol.* **2005**, *315*, 203–219. [CrossRef]
6. Olivera, F.; Choi, J.; Kim, D.; Li, M.-H. Estimation of average rainfall areal reduction factors in Texas using NEXRAD data. *J. Hydrol. Eng.* **2008**, *13*, 438–448. [CrossRef]
7. Overeem, A.; Buishand, T.A.; Holleman, I. Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. *Water Resour. Res.* **2009**, *45*. [CrossRef]
8. Overeem, A.; Buishand, T.A.; Holleman, I.; Uijlenhoet, R. Extreme value modeling of areal rainfall from weather radar. *Water Resour. Res.* **2010**, *46*. [CrossRef]
9. Cunnane, C. Methods and merits of regional flood frequency analysis. *J. Hydrol.* **1988**, *100*, 269–290. [CrossRef]
10. Burn, D.H. Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resour. Res.* **1990**, *26*, 2257–2265. [CrossRef]
11. Burn, D.H. An appraisal of the "region of influence" approach to flood frequency analysis. *Hydrol. Sci. J.* **1990**, *35*, 149–165. [CrossRef]
12. Pilona, P.J.; Adamowskib, K.; Alilab, Y. Regional analysis of annual maxima precipitation using L-moments. *Atmos. Res.* **1991**, *27*, 81–92. [CrossRef]
13. Svensson, C.; Jones, D.A. Review of rainfall frequency estimation methods. *J. Flood Risk Manag.* **2010**, *3*, 296–313. [CrossRef]
14. Hailegeorgis, T.T.; Thorolfsson, S.T.; Alfredsen, K. Regional frequency analysis of extreme precipitation with consideration of uncertainties to update IDF curves for the city of Trondheim. *J. Hydrol.* **2013**, *498*, 305–318. [CrossRef]
15. Naghavi, B.; Yu, F.X. Regional frequency analysis of extreme precipitation in Louisiana. *J. Hydraul. Eng.* **1995**, *121*, 819–827. [CrossRef]
16. Eldardiry, H.; Habib, E.; Zhang, Y. On the use of radar-based quantitative precipitation estimates for precipitation frequency analysis. *J. Hydrol.* **2015**, *531*, 441–453. [CrossRef]
17. Uboldi, F. A spatial bootstrap technique for parameter estimation of rainfall annual maxima distribution. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 981–995. [CrossRef]
18. Perica, S. *Precipitation Frequency Atlas of the United States, NOAA Atlas 14*; USA Department of Commerce, National Oceanic and Atmospheric Adminstration, National Weather Service: Silver Spring, MD, USA, 2013; Volume 9.

19. Smith, J.A.; Seo, D.J.; Baeck, M.L. An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.* **1996**, *32*, 2035–2045. [CrossRef]
20. Ciach, G.J.; Krajewski, W.F.; Villarini, G. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. *J. Hydrometeorol.* **2007**, *8*, 1325–1347. [CrossRef]
21. Villarini, G.; Krajewski, W.F.; Ciach, G.J.; Zimmerman, D.L. Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates. *Water Resour. Res.* **2009**, *45*, W01404. [CrossRef]
22. Habib, E. Independent assessment of incremental complexity in NWS multisensor precipitation estimator algorithms. *J. Hydrol. Eng.* **2013**, *18*, 143–155. [CrossRef]
23. Seo, D.J. Real-time estimation of rainfall fields using radar rainfall and rain gauge data. *J. Hydrol.* **1998**, *208*, 37–52. [CrossRef]
24. Kitzmiller, D.; Miller, D.; Fulton, R.; Ding, F. Radar and multisensor precipitation estimation techniques in National Weather Service hydrologic operations. *J. Hydrol. Eng.* **2013**, *18*, 133–142. [CrossRef]
25. Eldardiry, H.; Habib, E.; Zhang, Y.; Graschel, J. Artifacts in Stage IV NWS real-time multisensor precipitation estimates and impacts on identification of maximum series. *J. Hydrol. Eng.* **2015**, *22*, E4015003. [CrossRef]
26. Greene, D.R.; Hudlow, M.D. Hydrometeorologic grid mapping procedures. In Proceedings of the AWRA International Symposium on Hydrometeorology, Denver, CO, USA, 13–17 June 1982.
27. Faiers, G.E.; Keim, B.D.; Hirschboeck, K.K. A Synoptic evaluation of frequencies and intensities of extreme three-and 24-hour rainfall in Louisiana. *Prof. Geogr.* **1994**, *46*, 156–163. [CrossRef]
28. Madsen, H.; Rasmussen, P.F.; Rosbjerg, D. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resour. Res.* **1997**, *33*, 747–757. [CrossRef]
29. Hosking, J. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc.* **1990**, *52*, 105–124. [CrossRef]
30. Martins, E.S.; Stedinger, J.R. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour. Res.* **2000**, *36*, 737–744. [CrossRef]
31. Hershfield, D.M. *Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years*; Technical Paper 40; USA Weather Bureau: Washington, DC, USA, 1961.
32. Efron, B. Bootstrap methods: Another look at the Jackknife. *Ann. Stat.* **1979**, *7*, 1–26. [CrossRef]
33. Hosking, J.R.M.; Wallis, J.R. *Regional Frequency Analysis*; Cambridge University Press: Cambridge, UK, 1997.
34. Kachroo, R.K.; Mkhandi, S.H.; Parida, B.P. Flood frequency analysis of southern Africa: I. Delineation of homogeneous regions. *Hydrol. Sci. J.* **2000**, *45*, 437–447. [CrossRef]
35. Maddox, R.A.; Chappell, C.F.; Hoxit, L.R. Synoptic and meso-α scale aspects of flash flood events 1. *Bull. Am. Meteorol. Soc.* **1979**, *60*, 115–123. [CrossRef]
36. Habib, E. Ground based direct measurement. In *Rainfall: State of the Science*; Testik, F.Y., Mekonnen, G., Eds.; American Geophysical Union: Washington, DC, USA, 2010; Volume 191, pp. 61–77.
37. Weather Bureau. *Rainfall Intensity-Frequency Regime—Part 2: The Southeastern United States*; Technical Paper No. 29 (TP-29); USA Weather Bureau: Washington, DC, USA, 1958.
38. Vogel, R.M.; Shallcross, A.L. The moving blocks bootstrap versus parametric time series models. *Water Resour. Res.* **1996**, *32*, 1875–1882. [CrossRef]
39. Ciach, G.J.; Morrissey, M.L.; Krajewski, W.F. Conditional bias in radar rainfall estimation. *J. Appl. Meteorol.* **2000**, *39*, 1941–1946. [CrossRef]
40. Habib, E.; Qin, L. Application of a radar-rainfall uncertainty model to the NWS multi-sensor precipitation estimator products. *Meteorol. Appl.* **2013**, *20*, 276–286. [CrossRef]
41. Wazneh, H.; Chebana, F.; Ouarda, T.B. Delineation of homogeneous regions for regional frequency analysis using statistical depth function. *J. Hydrol.* **2015**, *521*, 232–244. [CrossRef]

Article

Spatio-Temporal Assessment of Global Precipitation Products over the Largest Agriculture Region in Pakistan

Zain Nawaz [1,2], Xin Li [3,4], Yingying Chen [3,4,*], Naima Nawaz [5], Rabia Gull [6] and Abdelrazek Elnashar [7]

[1] Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; zain-nawaz@lzb.ac.cn
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] National Tibetan Plateau Data Center, Key Laboratory of Tibetan Environmental Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China; xinli@itpcas.ac.cn
[4] CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing 100101, China
[5] Department of Rural Sociology, University of Agriculture, Faisalabad 38040, Pakistan; naimauaf@hotmail.com
[6] Department of Botany, University of Agriculture, Faisalabad 38040, Pakistan; Gull.rabi@gmail.com
[7] State Key Laboratory of Remote Sensing, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; abdelrazek.elnashar@cu.edu.eg
* Correspondence: chenyy@itpcas.ac.cn; Tel.: +86-10-84097114

Abstract: Spatial and temporal precipitation data acquisition is highly important for hydro-meteorological applications. Gridded precipitation products (GPPs) offer an opportunity to estimate precipitation at different time and resolution. Though, the products have numerous discrepancies that need to be evaluated against in-situ records. The present study is the first of its kind to highlight the performance evaluation of gauge based (GB) and satellite based (SB) GPPs at annual, winter, and summer monsoon scale by using multiple statistical approach during the period of 1979–2017 and 2003–2017, respectively. The result revealed that the temporal magnitude of all the GPPs was different and deviate up to 100–200 mm with overall spatial pattern of underestimation (GB product) and overestimation (SB product) from north to south gradient. The degree of accuracy of GB products with observed precipitation decreases with the increase in the magnitude of precipitation and vice versa for SB precipitation products. Furthermore, the observed precipitation revealed the positive trend with multiple turning points during the period 1979–2005. However, the gentle increase with no obvious break point has been detected during the period of 2005–2017. The large inter-annual variability and trends slope of the reference data series were well captured by Global Precipitation Climatology Centre (GPCC) and Tropical Rainfall Measuring Mission (TRMM) products and outperformed the relative GPPs in terms of higher R^2 values of ≥ 0.90 and lower values of estimated RME $\leq 25\%$ at annual and summer monsoon season. However, Climate Research Unit (CRU) performed better during winter estimates as compared with in-situ records. In view of significant error and discrepancies, regional correction factors for each GPPs were introduced that can be useful for future concerned projects over the study region. The study highlights the importance of evaluation by the careful selection of potential GPPs for the future hydro-climate studies over the similar regions like Punjab Province.

Keywords: gridded precipitation products; abrupt changes; trends; statistical indicators; agriculture; Pakistan

---

1. Introduction

Accurate and reliable estimates of global climate patterns are directly associated with the regional variation in precipitation [1]. The changes in amount and pattern of precipitation could directly influence the water resources and agriculture of the concerned regions [2]. Therefore, understanding the spatiotemporal variation in precipitation on the regional scales is of great importance in climate monitoring and in hydro-climate studies [3]. Several researchers have reported the spatiotemporal variations of precipitation for different regions of the world [4–7]. There is a growing agreement that long term changes in precipitation could alter the ecological and hydrological processes [8] and underpin our knowledge of global and regional climate change [9]. These accurate and reliable precipitation records underpin our knowledge of regional and global climate change, as well as their possible impacts on water resources [10,11].

In general, gauge measurements are the basic and reliable way of precipitation data acquisition [12]. Unfortunately, scarce gauge records, irregular distribution, limited data access, and poor spatial coverage hinder their use in conducting hydro-meteorological studies and climate change assessments [13,14]. In recent decades, with the advancement in remote sensing and geo-information technology, the gridded precipitation products (hereafter GPPs) has proven to be a reliable and cost-effective way of retrieving gridded precipitation data at various spatial and temporal scales across the globe [15]. These precipitation data either derived from satellite products or from the nationwide meteorological stations by using different interpolated algorithms and computational techniques by considering the physical characteristics (slope and elevation) of different regions. These multi-source data products are often applied as climatological input for hydro-climate simulation studies in data scarce extents to bridge the gap at regional scale [16] and there has been a considerable increase in the use of these products, owing to their easy accessibility, spatiotemporal coverage, and fine resolutions [17].

The evaluation of GPPs has also proved useful for different hydro-climate applications. The precipitation variability in different Gridded Data Products (GDPs) has been quantified for different regions across the globe [18–20]. Various studies were carried out in recent years to assess the performance of GDPs, revealing considerable differences between the products at the regional scale [21]. Furthermore, there are uncertainties that are associated with GPPs, because of the variability in spatial and temporal coverage, lack of in-situ observations, relocation of gauges, and data processing practices [22]. Thus, the reliability and accuracy of GPPs varies with time and regional climate [23]. Therefore, it is highly important to assess and evaluate the performance and capability of the GPPs at regional scales, especially the arid and semi-arid regions that are more sensitive to insignificant changes in climatic characteristics due to its fragile ecosystems [24]. Such regions are characterized by very complex hydrological systems that often exhibit extreme behaviors, such as extended droughts due to prolonged dry spell and floods due to high-intensity precipitation [25]. The predominantly arid and semi-arid climate and geographical location in the fast temperature rising region have made Pakistan one of the most vulnerable countries in the world to climate change [1]. Moreover, the natives of the country are mostly engaged in agriculture, a highly susceptible sector to climate change, with limited resources to adapt to changing circumstances [26].

Few studies were carried out in order to evaluate the performance of different GPPs against the reference data, which mainly focused on the basin level or higher altitude sites of the country and ignored the important segment of spatiotemporal precipitation variations in agricultural region of the country. For instance, [27] assessed the precipitation distribution in the high altitude region of Hindu-Kush-Himalaya basins by using different precipitation products. The study reported the better performance of ECMWF Re-Analysis (ERA-Interim) product at high catchments as compared with WATCH Forcing Data Methodology (WFDEI) and APHRODITE products. Using different gridded precipitation products, [28,29] reported the significant errors in different gridded product; however, Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) product performed better in the high catchments of the Indus basin. On the contrary, [30] reported

the overestimation of TMPA products against reference data using the simple statistical metric approach over the complex topography of Pakistan. The present study is first of its kind, which aims to bridge the gap of knowledge with a detailed multiple-scale assessment of the spatiotemporal uncertainties of selected global precipitation products that are generated by different sources by using different statistical metrics, trends evaluation and comparison approach over the Punjab province, Pakistan. It is worth mentioning that Punjab province is highly important in the perspective of agriculture and irrigated farming as it produces major agriculture commodities of the country and it is highly vulnerable to changes in most of meteorological parameters with high frequency events and, hence. is highly prone to climate change [31]. In this study, we aim to assess the quality and differences of the different GPPs generated from multiple sources, i.e., Gauge Based (GB) products, Global Precipitation Climatology Centre (GPCC), Center for Climatic Research, University of Delaware (UDel), Asian Precipitation—Highly-Resolved Observational Data Integration Towards Evaluation (APHRODITE), Climate Prediction Centre (CPC), Climatic Research Unit, University of East Angelia (CRU) and Satellite Based (SB) products, Tropical Rainfall Measuring Mission (TRMM, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks- Cloud Classification System (PERSIANN-CCS), Global Satellite Mapping of Precipitation (GSMap) and Climate Hazard Group Infrared Precipitation with Station Data(CHIRPS) during the period of 1979–2017 and 2003–2017, respectively, over the Punjab province in Pakistan. Secondly, to evaluate and compare the changes in temporal trends and abrupt turning points in selected GPPs against the reference data for the study region. We perceive the usefulness of this study as multi-directional, because the findings of the study could be used as baseline for the selection of potential GB and SB GPPs over multiple time scales for different hydro-meteorological studies.

2. Study Area

Pakistan is geographically located in southwest Asia with an area of 8×10^6 km^2 between latitude and longitude range of 24–37° N and 60–75° E (Figure 1). The country has diverse topography that ranges from Karakoram and Himalayan mountains in the north and northwest to the agriculture plains in the center and south of the Indus basin along the southern cost of the Arabian sea [32]. Pakistan is an agrarian country and Punjab is the second largest Province with geographical coordinates of 31.17° N and 72.70° E. The Province has the largest population and producing more than 50% of the country's agricultural commodities [33]. The regional mean temperature varies from 23 to 26 °C, with a T_{min} of 16–19 °C and a T_{max} of 29–33 °C. The northern part of the Province receiving more precipitation than southern part with overall annual mean precipitation ranges from < 300 mm in the southern part to > 800 mm in the northern part [34]. The two major seasons that dominates the overall hydrology of the region, i.e., the summer monsoon (June–September) and winter monsoon seasons (December–April) [35]. The maximum precipitation occurs during the summer monsoon due to the monsoon system that originates from Bay of Bengal and enters in the country from east and northeast [35]. The winter precipitation instigates from the Mediterranean Sea due to western disturbances and enters in the country from southwest and northeast direction [35]. There is great variation and extreme events during summer and winter monsoon precipitation seasons that make this region a highly susceptible to climate change [36]. Due to highly vulnerable in precipitation changes, the spatial and temporal evaluation of the GPPs is highly important for further application of these products, including hydro-meteorological studies over the study region.
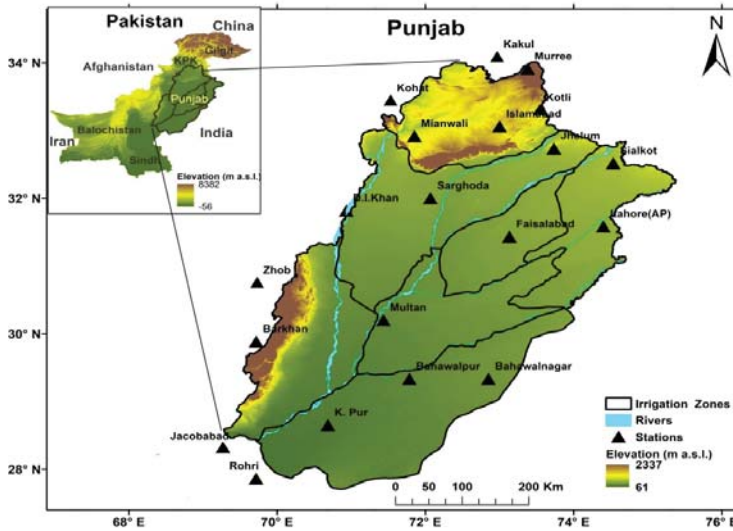
Figure 1. Study area and meteorological stations.

3. Materials and Methods

3.1. In-Situ Records

The long term monthly precipitation data that were collected from 20 meteorological stations over the Punjab region were acquired from Pakistan Meteorological Department (PMD) (Figure 1). The in-situ records were investigated during the study periods from 1979 to 2017 and from 2003 to 2017 as reference data for the assessment of GB and SB precipitation products, respectively. The selection of the gauges depends on their maximum availability and completeness of data series. The GPPs were evaluated against the in-situ records on an annual, winter, and summer monsoon season. The winter and summer monsoon seasons are the primary drivers of the annual hydrological cycle with major influence of Asian monsoon (summer monsoon season) over the study region [28]. The GB and SB GPPs were developed while using quality-controlled procedures and algorithms. However, the stations data are generally considered as standard measurements for the evaluation of GPPs [37,38]. Moreover, the standard quality-controlled procedures were adopted for the better accuracy assessment of GB and SB GPPs. For quality assurance, station data series were thoroughly checked and outliers were fixed with neighboring gauge records [39]. Gaps were filled using a time-based interpolation approach [40]. Several post-processing techniques have been developed to identify the consistency and homogeneity's in the station records [41]. The double-mass curve was applied to observe the inhomogeneity in the station records [42]. The curve of all the stations showed a straight line with no evident break points, which confirms a temporal consistency and uniformity in the time series records.

After quality control process in the reference data series, the thin-plate splines (TPS) interpolation method was introduced to convert the station data into products relative spatial grid sizes for the assessment of GB and SB GPPs on spatial scale. The original TPS method is defined by [43], whereas [44] provides a detailed description of its application for different climate indicators. The TPS scheme is applicable and robust in regions where stations density is low and prior estimation of the spatial auto-covariance is not required [44]. The time series autocorrelation method was applied prior to estimate the trend significance, magnitude, and abrupt turning point to ensure the occurrence of significant autocorrelation in the reference data series [1]. The presence of significant autocorrelation in the reference data series could influence the outcome of Mann–Kendall (MK), Sen's Slope (SS),

and Mann–Kendall abrupt change analysis [45]. Therefore, the significance of autocorrelation should be checked before applying the trend tests [46]. The analysis revealed no significant autocorrelation in the reference data series of annual, winter, and summer monsoon seasons. Therefore, the in-situ records are completely independent, and the MK test is applicable to the original gauge records. The detailed description for autocorrelation method is reported by [47].

3.2. Gridded Precipitation Products (GPPs)

In this study, GB and SB GPPs (GPCC, UDel, APHRODITE, CPC, CRU, TRMM, PERSIANN-CDR, PERSIANN-CCS, and CHIRPS, GSMap) were evaluated against the reference data (Table 1). Details of the datasets are described below.

3.2.1. Gauge-Based GPPs

The Global Precipitation Climatology Centre (GPCC) obtained the primary data from meteorological agencies at National level (NMAs), Food and Agriculture Organization (FAO), Climate Research Unit, University of East Angelia (CRU), and Global historical climatology network (GHCN) at the National Centers for Environmental Information [48]. The product covers the period from 1891 to 2018. In this research, the monthly product of 0.5° spatial resolution is used from the period 1979–2017. The product from University of Delaware, USA (UDel) acquired the primary data from GHCN2, daily GHCN from National Centers for Environmental Information, National Center Atmospheric Research (NCAR), Project Greenland of automatic meteorological stations, Data archives of Nicholson for African continent, records of daily summary at the global level [49]. The product covers the period from 1900 to 2017. The monthly precipitation data of 0.5° spatial resolution from the period 1979–2017 is utilized in the study. The monthly product Climate Research Unit (CRU) acquired the primary data from the World Meteorological Organization (WMO), Food and Agriculture Organization (FAO), and National meteorological departments. The CRU product covers the period from 1901 to 2019. The monthly precipitation data of 0.5° spatial resolution from the period 1979–2017 is used in the present study [50].

The daily Asian Precipitation-Highly Resolved Observational Data Integration towards Evaluation of Water Resources (APHRODITE) product that was obtained the primary source data from multiple Asian agencies including Global Telecommunication System (GTS), National weather agencies, International Centre for Integrated Mountain Development (ICIMOD), International Water Management Institute (IWMI) and other national and international projects on climate [51]. The product covers the period from 1961 to 2015. The daily precipitation product of 0.5° spatial resolution was accumulated into monthly time scale from the period 1979–2015 in the present study. The CPC is the first product of Unified Precipitation Project at the National Oceanic and Atmospheric Administration (NOAA). The product acquired the primary data from National weather agencies, quality controlled station data from GTS, and Cooperative Observer Network [52]. The CPC product covers the period from 1948 to 2018. The daily precipitation product of 0.5° spatial resolution was accumulated into monthly time scale from the period 1979–2017 in the present study.

3.2.2. Satellite-Based GPPs

The TRMM satellite provides the continuous precipitation data (1998–present) which covers the range from 50° N to 50° S over tropical to subtropical regions at 0.25° spatial resolution [53]. The TRMM Multi-satellite Precipitation Analysis (TMPA) used three basic instruments to record the data, including Visible Infrared Scanner, Microwave Image and Radar Precipitation [30]. The TMPA provides real time and post real time products. The post real time version 7 products have adjusted with gauge data and proved higher accuracy than TMPA real time product [54]. The present study utilized the higher accurate version 7 products with monthly temporal and $0.25° \times 0.25°$ spatial resolution during the study period 2003–2017.

The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) provides longer gridded data (1983–present) with spatial resolution of 0.25° × 0.25° at the daily timescale. The PERSIANN-CDR estimates the rainfall rate from geostationary satellites by using infrared brightness temperature [55]. The stage IV radar data from the National Centers for Environmental Prediction (NCEP) is used to train the Artificial Neural Network (ANN) model. The high resolution precipitation estimates then attuned by Global Precipitation Climatology project (GPCP) for bias correction [56]. The daily data aggregated into monthly time scale for evaluation against reference data for the study period 2003–2017.

The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System (PERSIANN-CCS) provides the gridded precipitation data (2003–present) at spatial resolution of (0.04° and 30 min.) by using Infrared brightness temperature derived from geostationary satellites and continuously updating its parameter using Passive Microwave (PMW) measurements from low earth orbit satellites. The regression and histogram matching are used to draw fit curve plot between temperature brightness of pixel and rainfall rate to achieve rainfall mapping of each classified cloud cluster [57]. The daily data aggregated into monthly time scale for evaluation against reference data for the current study period 2003–2017. The CHIRPS precipitation dataset is quasi land product belongs to Climate Hazard Group with a spatial resolution of 0.05 °s and temporal daily scale derived from TRMM satellite, and several observed products such as NOAA, CPC, National Climatic Data Center (NCDC) and Climate Forecast System version 2 (CFSv2). The product algorithm is based on the concept of cold cloud duration (CCD), which is the duration of time of pixel covered by IR brightness temperature. Precipitation estimates by using CCDs procedure by incorporating TMPA3B42 product and merged with observed measurements using the Inverse Distance Weighting (IDW) method to produce final product [58]. For the current study, the daily data aggregated into monthly time scale during the study period 2003–2017. The recent GSMap project is launched by Japan Aerospace Exploration Agency (JAXA) to monitor precipitation at higher spatial resolution of 0.1° and one-hour temporal scale. The product input is based on multiple polar orbiting satellites with adjusted accuracy of Kalman smoothing approach [59]. The product covers the range of latitude and longitude from 60° N to 60° S and from 180° W to 180° E, respectively. For the present study, the hourly data converted into monthly time series during the period 2003–2017.

Table 1. Information of gauge based (GB) and satellite based gridded data products (SB GPPs) used in the study.

Datasets	Resolution	Frequency	Coverage	Study Period	Reference
APHRODITE	0.5°	Monthly	Global land	1979–2015	[51]
CRU	0.5°	Monthly	Global land	1979–2015	[50]
CPC	0.5°	Monthly	Global land	1979–2015	[52]
GPCC	0.5°	Monthly	Global land	1979–2013	[48]
UDel	0.5°	Monthly	Global land	1979–2015	[49]
TRMM	0.25°	Monthly	Global land	2003–2017	[53]
GSMap	0.1°	Monthly	Global land	2003–2017	[59]
CHIRPS	0.05°	Monthly	Global land	2003–2017	[58]
PERSIANN-CDR	0.25°	Monthly	Global land	2003–2017	[55]
PERSIANN-CCS	0.04°	Monthly	Global land	2003–2017	[57]

3.3. Descriptive Methods

Here, we assess the spatial and temporal performance of each GPP against the in-situ records using different statistical techniques. The quantitative statistical indicators, including the root mean square error (RMSE), Pearson correlation coefficient (CC), and standard deviation were applied while using a Taylor diagram, which is a precise technique of measuring the degree of accuracy between GPPs and reference data [60]. The mean error (ME) and relative mean error (RME) were calculated in order to indicate the systematic bias which determines the level of over-or under-estimation of

GPPs against in-situ data. Furthermore, the scatter plots were also used to determine the quantitative linear relationship between each GPP against the reference data. Furthermore, the detection and comparison of trends in different GPPs data series were evaluated against the reference data using the non-parametric Mann–Kendall (MK) test. The MK trend test is simple and it has been widely used for the detection of significant trends in hydro-meteorological time series data [61,62]. The trend test is robust against normal distribution, missing values, outliers and is less susceptible to the abrupt change point [63].

Similarly, the Theil and Sen's Slope (TSS) is non-parametric test, which can be used to quantify the slope magnitude in linear trends [64]. The TSS has been widely acceptable and used by many researchers to detect the significant trends in different climate indicators [19]. The test is based on least square regression technique, which is commonly used to estimate the rate of slope in a given time series data [65,66]. Moreover, the abrupt change analysis was performed using the Sequential Mann Kendall (SQMK) test. The SQMK test was proposed by [67] and it has been widely used to identify the abrupt change point in hydro-meteorological time series data [49]. The test sets up two temporal series based on forward and backward process i.e., progressive series (PS) and retrograde series (RS). In this test, the progressive series is a standardized variable with zero mean and unit standard deviation. The nature of the progressive series is same as that of Z values, which range from the initial to last data point. Similarly, the value of the retrograde series is computed backwards, starting from the end point and finishing at the first point of the temporal series. The positive and negative change in time series data indicate increasing and decreasing trends, respectively [67]. Details of the descriptive statistical methods are discussed and reported by many research papers [68–71]

4. Results

4.1. Evaluation of Temporal and Spatial Dynamics

Annual and Seasonal Scale

The annual averages of precipitation, as estimated from reference and gridded data products (GPPs) over Punjab region, are shown in Figure 2. The temporal trends indicated that the annual precipitation amounts that were underestimated by the GPPs with the exception of TRMM and PERSIANN-CDR products, which overestimated the magnitude as compared with the reference data. The temporal magnitude in all the precipitation products was different and deviate up to 100–200 mm as compared with in-situ measurements. However, the most identical temporal fluctuation pattern of precipitation was observed in GPCC and TRMM products with the relative mean error of −16.59% and 2.88%, respectively. Figure 3 illustrates the spatial distribution of GPPs and observed annual average precipitation. The results indicated that, the spatial configuration of GPPs exhibited the similar pattern of precipitation as compared with observed estimates. Across all of the GPPs, the northern part of study area received more amount of precipitation than southern part and exhibited the precipitation pattern towards south to north gradient with the range of 15–1500 mm. In all GPPs, the spatial consistency was observed with a major deviation in northern Punjab. However, the most accurate and consistent spatial pattern was observed in GPCC and TRMM product and the lowest spatial accuracy was observed in CPC following by UDel and PERSIANN-CCS products particularly in northern Punjab. These products were relatively less accurate to capture the higher amount of precipitation distribution over the northern part of the study region. In contrast, the APHRODITE product showed less precipitation magnitude in the central and southern Punjab as compared with observed precipitation. The spatial distribution of the APHRODITE product showed better results at high catchments of the study region as compared with plain areas. These results are consistent with the findings of [27], who reported the better performance of APHRODITE product as compared with reference data at high-altitude catchments in Pakistan.
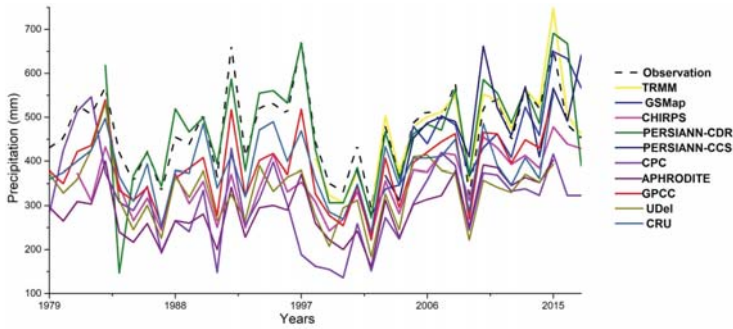
Figure 2. Mean annual variation in GPPs and reference data.
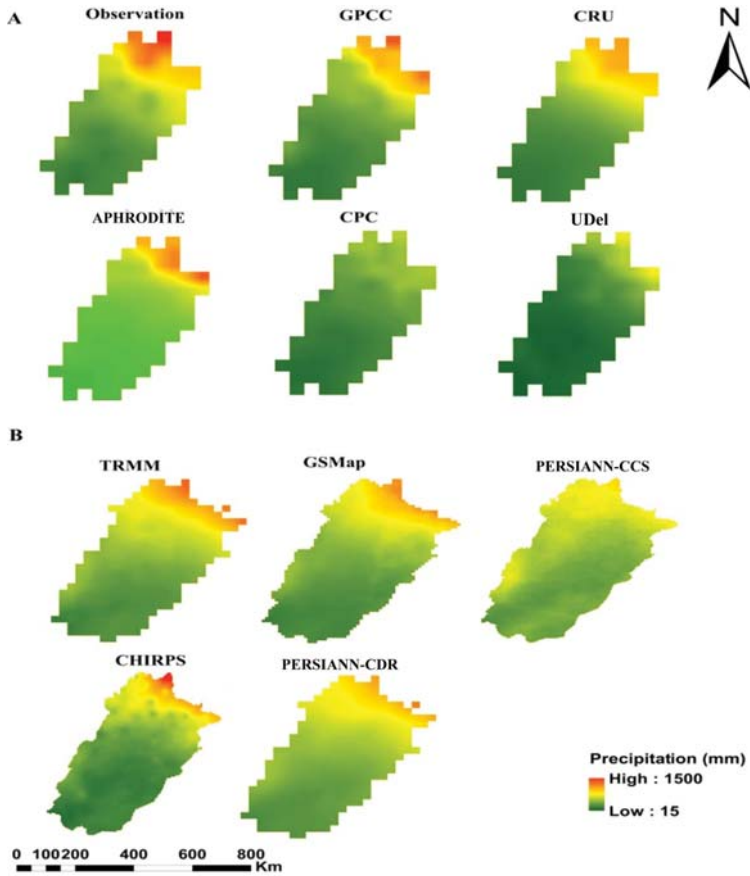
Figure 3. Spatial distribution of annual average of (**A**) GB and (**B**) SB precipitation products during the whole study periods.

The descriptive statistical measures divulge the characteristics of GPPs when compared with the reference data. Figures 4–6 indicates the plots as A (GB) and B (SB) precipitation products against the reference data during annual, winter, and summer monsoon, respectively. The result showed that, all the selected GPPs underestimated the annual precipitation amount, except TRMM and PERSIANN-CDR

product, which indicates the overestimation of precipitation over the study region. The highest accuracy was observed in the GPCC and TRMM products with lowest mean error and higher R-squared value with the magnitude of −76.35 (0.90) and 14.26 mm (0.90) respectively. These products are comparatively better to capture the annual precipitation variability over the study region. However, the lowest accuracy was observed in CPC and PERSIANN-CCS products with the magnitude of mean error −163.57 (−34.71%) and −16.08 mm (−3.23%), respectively. These datasets are relatively less accurate for capturing the annual variability amount of precipitation during the whole study period. Moreover, the GSMap and PERSIANN-CCS products overestimated the precipitation magnitude ≤ 500 mm and underestimated the precipitation amounts, which are ≥ 500 mm. The inter-comparison of gridded products indicated a similar consistency, except TRMM and PERSIANN-CDR products, which showed positive bias. The underestimation was more conspicuous in GB products as compared with SB products. The SB products showed lower bias and higher accuracy as compared with GB products, which could be due to higher pixel resolution of SB products over the study region [28]. Overall, TRMM product indicated better quantitative performance and showed reasonable consistency against reference precipitation over the study region.
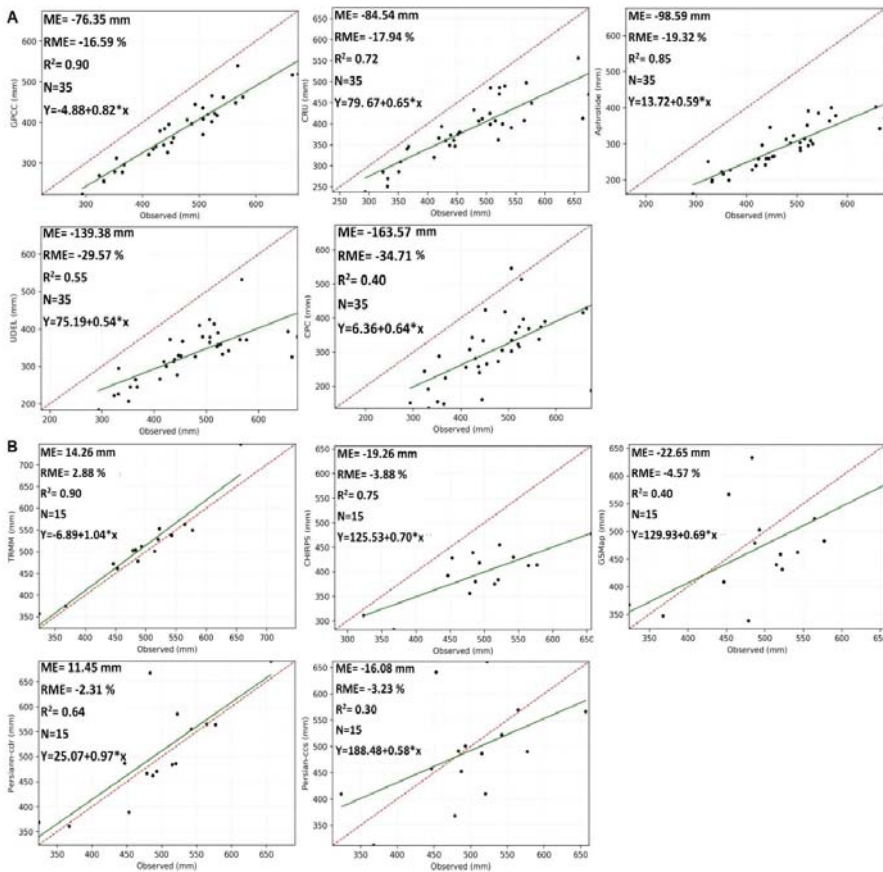
Figure 4. Statistical indicators for the assessment of (**A**) (GB) and (**B**) (SB) precipitation products against reference data (Annual timescale).

The GB and SB precipitation products were further evaluated during winter monsoon season. All of the GB products underestimated the winter precipitation with the highest and lowest accuracy being observed in CRU and CPC product with the estimated mean error of −25.08 (−20.49%) and 37.99 mm (−31.04%), respectively. In contrast, the SB products overestimated the precipitation magnitude except CHIRPS product, which indicates the underestimation of precipitation amount over the study region. The highest and lowest accuracy was observed in TRMM and GSMap products with the estimated mean error of 24.06 (19.39%) and 80.66 mm (65.03%) respectively. The rest of the precipitation products also showed better quantitative agreement with a higher range of deviation from the mean observed precipitation. Overall, the performance of CRU and TRMM were consistent and show relatively less error during the winter monsoon season.
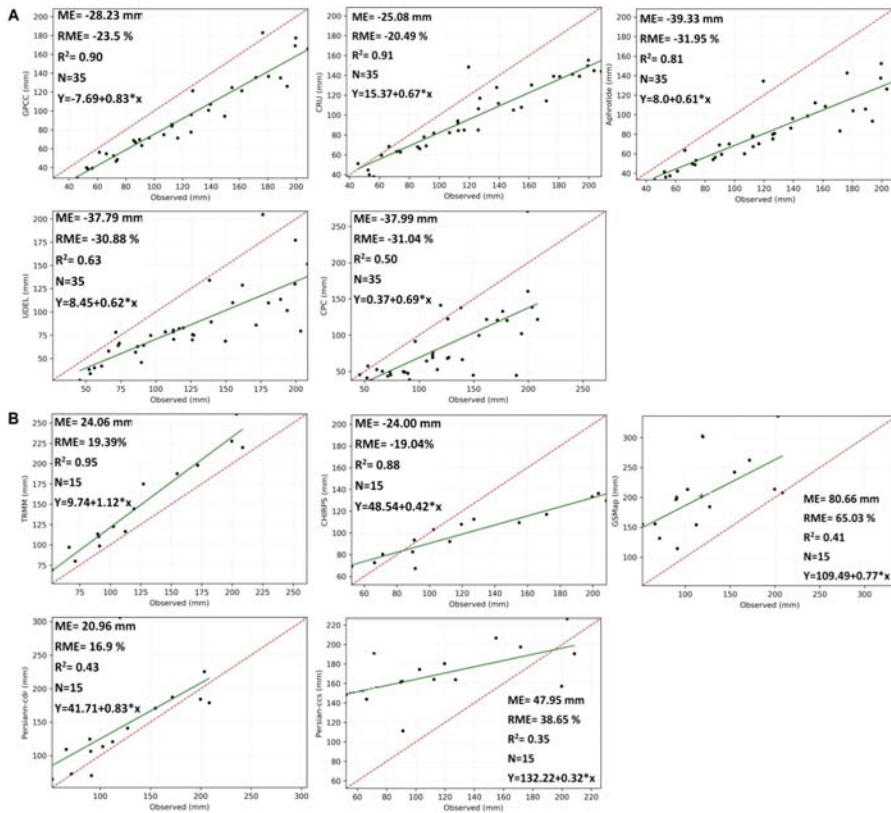
Figure 5. Statistical indicators for the assessment of (**A**) (GB) and (**B**) (SB) precipitation products against reference data (winter monsoon).

Furthermore, the evaluation of precipitation products during the summer monsoon season indicated the similar pattern of statistical metrics as assessed during annual time scale. The summer monsoon precipitation dominated the regional water balance and plays a significant role in annual precipitation variability. All of the GB and SB products underestimated the observed precipitation. The highest and lowest accuracy in GB products was observed in GPCC and CPC with the estimated mean error of −50.03 (−16.17%) and −114.50 (−36.48%), respectively. Moreover, the TRMM and GSMap showed the highest and lowest accuracy as compared with in-situ with the estimated mean error of −15.18 (−4.48%) and −136.06 mm (−40.16%), respectively. The products underestimation was more

conspicuous during winter monsoon period as compared with annual and summer monsoon season. Moreover, all of the SB products showed less range of mean error during annual scale following by summer and winter monsoon seasons. However, the GB products exhibited lower mean error in winter following by annual and summer scale. Overall, the performance of GPCC and TRMM products showed the best agreement with in-situ in terms of higher R^2 values of ≥ 0.90 and lower values of RME $\leq 25\%$ during annual, winter, and summer monsoon seasons.
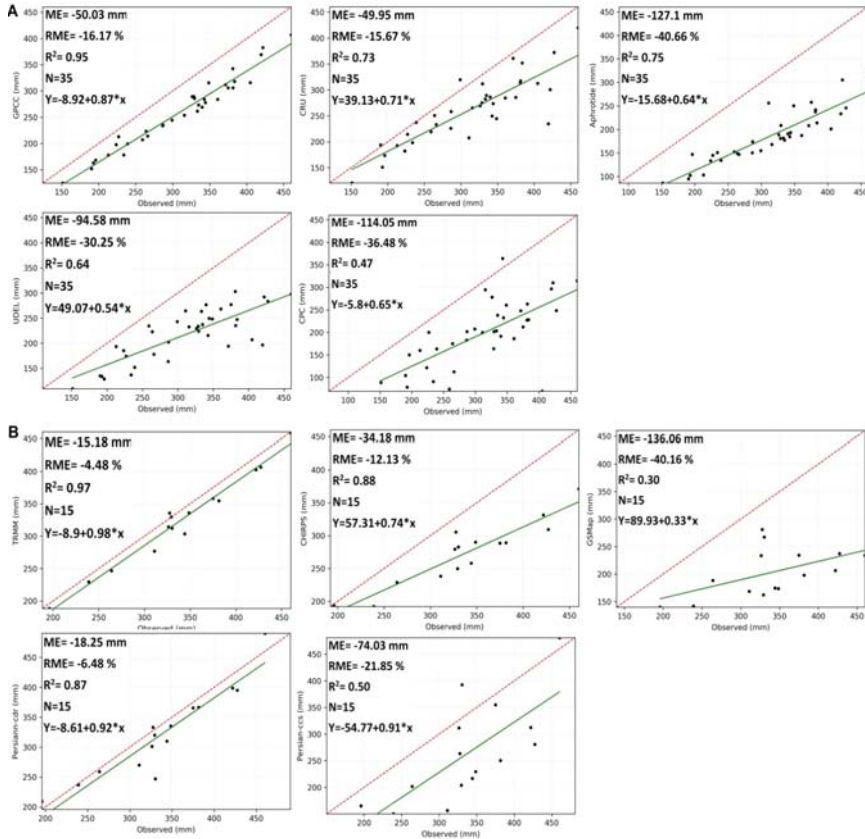
Figure 6. Statistical indicators for the assessment of (**A**) (GB) and (**B**) (SB) precipitation products against reference data (summer monsoon).

In order to evaluate the temporal pattern variability of different GPPs against observation data, Taylor diagrams were plotted [62] to quantify the precise agreement between the observation and GPPs in terms of correlation coefficients (CC), standard deviation (SD), and root mean square deviation (RMSD), which are shown in Figure 7 as A (GB) and B (SB) precipitation products during annual, winter, and summer seasons. In the diagram, correlation coefficient (CC) is denoted by blue lines adjoining perpendicular to the parabolic scale, standard deviation (SD) is denoted by radii of the black cycles and root mean square deviation (RMSD) donated by the radii of the green cycles. In the diagram, if the value of the GPPs is closer to the observation data, then it is considered to be a better product. The diagram statistics provide the evaluation of the temporal pattern of the GPPs against observed data.
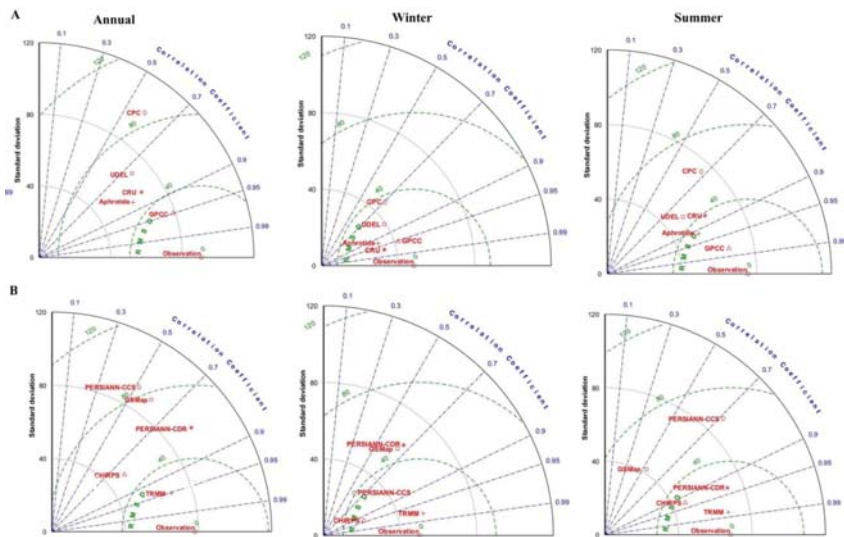
Figure 7. Statistical evaluation of (**A**) (GB) and (**B**) (SB) precipitation products against reference data by using Taylor diagram.

The results indicated that the GPCC and TRMM precipitation products outperformed the other products during the annual scale in terms of higher CC of around 95% and less value of the RMSD. The GPCC and TRMM points marked closer to the reference point indicated that the products relatively performed better and they are suitable for the study region. The rest of the precipitation products also depicted good agreement with observed data over the study region, with notable mean error all the way. These results are in agreement with the findings of [27,72] who reported the better performances of TRMM and GPCC products. The range of agreement is different, which could be due to a difference in region, study period and statistical metrics. Moreover, the Taylor analysis during winter precipitation indicated the better performance of CRU and TRMM products as marked closer to the reference data with relatively higher CC of around 97% and lower values of RMSD. The GPCC, APHRODITE, and CHIRPS products also showed better efficiency in terms of high CC and lower RMSD values. On the other hand, the evaluation of precipitation products during summer monsoon revealed the better performance of GPCC and TRMM products, as indicated by a higher CC of around 98% and lower RMSD values. These results are consistent with the outcomes of products evaluation during the annual time scale over the study region. The efficiency of GPCC and TRMM precipitation products during annual and summer season was found to be similar and consistent. The major influence of annual precipitation variability depends upon the summer monsoon season, as it receives the major amount of precipitation and main driver of annual hydrological cycle [73]. Overall, the performance of GPCC and TRMM products were best during the annual and summer time scale over the study region. However, CRU product shows relatively better performance during the winter monsoon period.

Figure 8 shows the spatial distribution of GB and SB precipitation products against reference data in terms of spatial pattern of statistical indices (ME, RMSD and CC) over the Punjab region during the whole study periods. The results indicated that all of the GB products underestimated the precipitation amounts as compared with observed precipitation. The range of deviation was more conspicuous in northern Punjab with the dominated pattern of underestimation from north to south gradient. In the entire study period, the GPCC product exhibited best agreement, as indicated by the distribution of higher CC and lower values of ME and RMSD. However, the accuracy of GPCC product was more precise over northeast and southern part of Punjab. The inter-comparison of GB products indicated the similar pattern of underestimation, except the UDel and CPC product, which indicated the higher

range of underestimation over the whole study region. Both of the products showed the lowest agreement following by APHRODITE product as compared with the observed pattern of precipitation. On the other hand, the SB precipitation products showed the visible pattern of overestimation over the whole study region. However, the pattern was more conspicuous over the northern and eastern part of the study region. Moreover, all of the products showed a similar pattern of underestimation over south-west part of the study region. However, the pattern of underestimation was more noticeable in PERSIANN products over south to south west gradient as compared with observed precipitation. The statistical pattern of all the SB products pointed out the better performance of TRMM product in terms of higher CC and lower ME and RMSD values. The TRMM product outperformed the rest of the product by showing the lowest error and highest accuracy. However, the pattern of overestimation was more noticeable over the northern region of the study area. The overestimation in these areas could be related to the higher rate of evaporation. These results are consistent with the findings of [27], who reported the higher overestimation of satellite products in the foothills of northern mountain range of Hindukush, Karakoram and Himalaya (HKH in Pakistan. Generally, the GPCC and TRMM products showed the better accuracy and consistency in terms of spatial distribution of statistical metrics with higher values of CC and lower values of ME and RMSD.
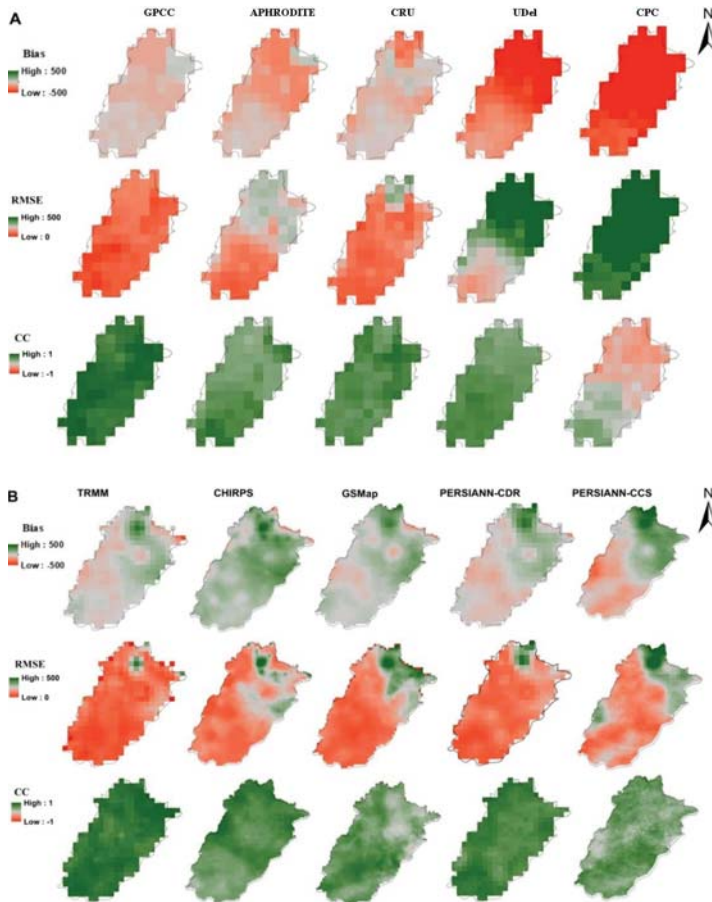
Figure 8. Spatial statistical indicators for the assessment of (**A**) (GB) and (**B**) (SB) precipitation products against reference data.

Overall, the extent of under/over estimation of precipitation by GB and SB products suggests the significance of data correction before their reliable utilization in climate and hydrological studies. For the correction of precipitation products, we analyzed the estimated mean error in different data products and introduce a correction factor of each GB and SB product for annual, winter, and summer monsoon season over the Punjab region. The regional correction factor of each product for annual, winter, and summer monsoon during the period of 1979–2017 and 2003–2017 for GB and SB precipitation products are summarized in Table 2. The correction factor needs to be multiplied with the respective data products in order to minimize the percentage of error over the target region. The corrected product data could be directly used in climate modelling and other relevant studies [27].

Table 2. Annual and seasonal correction factors for each GPPs over Punjab region.

GPPs	Annual	Winter-Monsoon	Summer-Monsoon
CRU	1.19	1.24	1.17
GPCC	1.18	1.31	1.15
APHRODITE	1.61	1.48	1.71
CPC	1.53	1.47	1.57
UDel	1.20	1.33	1.38
TRMM	0.99	0.84	1.06
CHIRPS	1.34	1.37	1.3
GSMap	1.05	0.61	2.33
PERSIANN-CDR	1.01	0.92	1.08
PERSIANN-CCS	1.01	0.72	1.79

4.2. Evaluation of Changing Trends and Abrupt Transition:

Figure 9 presents the annual trends of precipitation products and observed data for Punjab region obtained by the MK and Theil–Sen approach. The trend plots are arranged as A and B with respect to GB and SB products during the study periods 1979–2017 and 2003–2017, respectively. The results indicated that, the annual observed precipitation revealed the insignificant positive trend during the periods 1979–2017 and 2003–2017 with the rate of 1.12 and 5.5 mm/decade, respectively. All of the GB and SB products exhibited the non-significant positive trend with the different scale of overestimation, except for CRU, CPC, and UDel products, which indicates the insignificant trend with the magnitude of 0.05, −1.38, and −0.06 mm/decade, respectively. These products are relatively less accurate to detect the observed precipitation trend. Though, GPCC and TRMM products outperformed other datasets in terms of trends magnitude with the rate of 1.29 and 5.53 mm/decade, which is quite close to the magnitude of observed precipitation. The range of overestimation was higher in SB products, which could be due to the shorter time period. The highest increasing trend was indicated by APHRODITE product followed by GSMap and PERSIANN products as compared with the observed precipitation trend. Overall, the performance of GPCC and TRMM precipitation products showed better agreement in terms of trend detection when compared with the observed precipitation.
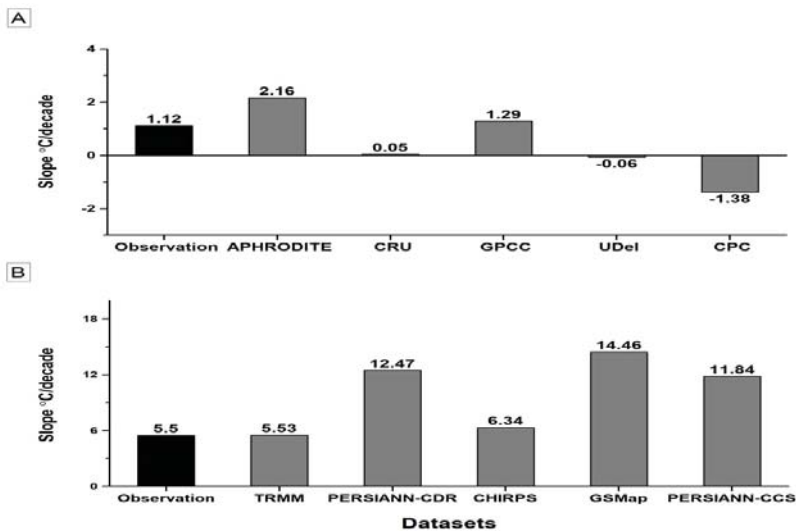
Figure 9. Comparative trend assessment of (**A**) (GB) and (**B**) (SB) precipitation products and reference data by using MK test (95% confidence Interval).

The abrupt changes in climate data series revealed the transition of climate from one state to another due to some external factors, which activates a change to new state at a rate determined by the climate system [74,75]. The Sequential Mann-Kendall (SQMK) test was applied for the evaluation and comparison of different precipitation products with observed data over the Punjab region. The retrograde and progressive series were attained at 0.05 significance level. The comparative evaluation of GB and SB precipitation products with reference data using the SQMK test with the confidence interval of 95% during the study periods 1979–2017 and 2003–2017 over the Punjab region are shown in Figure 10. The results indicate that the GPCC and TRMM products exhibited a similar pattern and mutation points as compared with reference data series. However, the other products also showed the similar fluctuated pattern in progressive series as compared with the reference data, yet they failed to capture the exact temporal mutation changes. The major fluctuations in the reference data series were detected during the period of 1984–1985 (negative), 1988–1989 (positive), 1998–1999 (negative), and 2005–2006 (positive). Moreover, the reference data showed the multiple turning points in progressive data series during the whole study Period. However, the gentle increase in the precipitation trend has been observed with no significant breakpoint during the period 2005–2017. All of the GB and SB products indicated the similar pattern of variability in progressive series. However, the TRMM product showed the best agreement as compared with the reference data series during this period over the Punjab region. The inter-annual and multi-decadal temporal abrupt changes in precipitation data series depicted a large fluctuation in precipitation over the targeted region during the whole study period. The major variabilities in the observed data series are well covenant with the findings of [71,76], who reported the similar variability in precipitation over India and Pakistan, respectively. Overall, the GPCC and TRMM precipitation products outperformed the other products and revealed the best agreement when compared with reference data to capture the temporal pattern and mutation points in the precipitation data series over the target region.
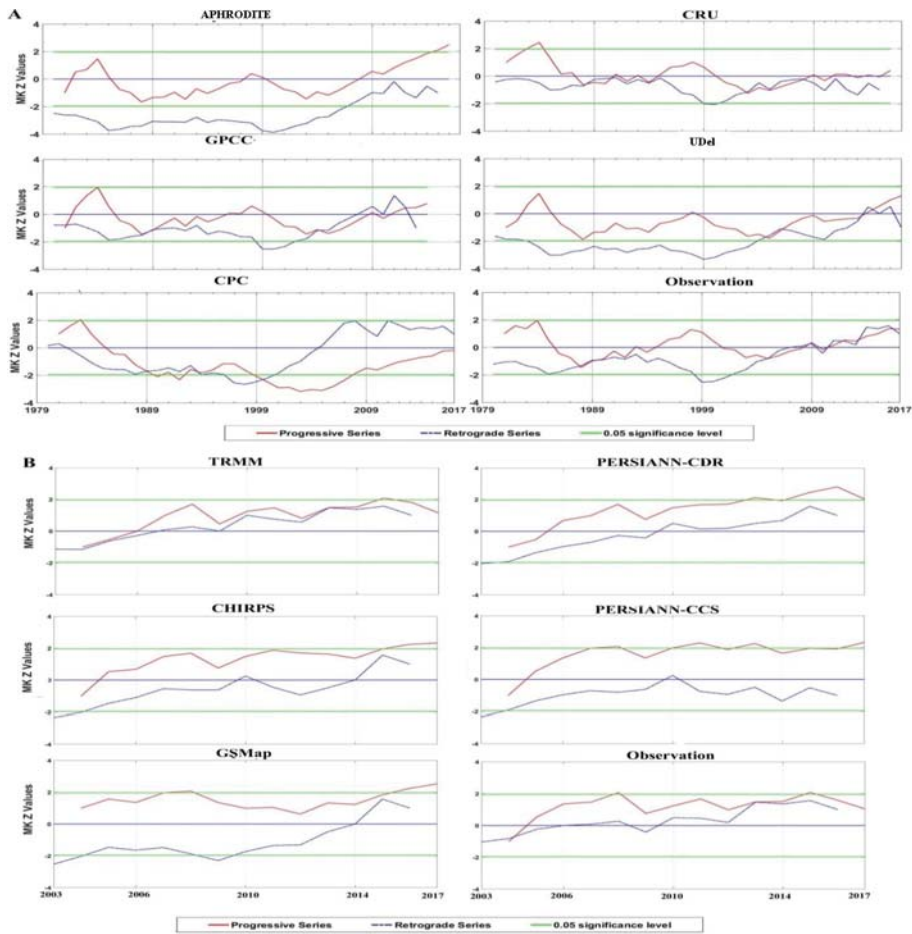
Figure 10. Comparison of Abrupt change detection in (**A**) (GB) and (**B**) (SB) precipitation products and reference data by using SQMK test (95% confidence Interval).

5. Discussion

Spatial and temporal precipitation data acquisition is highly important for the regional climate studies as well as for the management of water resources and agriculture of the concerned regions. With the advancement in geo-information and remote sensing technologies, this paves the way to acquire gridded data free of cost on a different time and scale. However, the reliability and accuracy of the data must be quantified against the in-situ records at the regional scale [77]. The current study provides a comprehensive assessment of GB and SB precipitation products against the reference records over the Punjab province during annual, winter, and summer monsoon from the period of 1979–2017 and 2003–2017, respectively. Moreover, the changes in trends and detection of abrupt turning point in temporal data series of products were evaluated and compared with the reference data series. In addition, the correction factor for each product was introduced to minimize the percentage of error in the respective products.

The results indicate that the spatial and temporal performance of GPCC and TRMM precipitation product outperformed the other corresponding products in terms of statistical measurements (high value of CC and R^2 with lower value of ME and RMSD) during the annual and summer monsoon season

during 1979–2017 and 2003–2017, respectively. However, CRU performed better during winter estimates when compared with reference data during the whole study period. Moreover, the range of deviation among different precipitation products underestimated with different magnitude as compared with the reference data. We found large uncertainties in the magnitude and temporal variability among different products. The magnitude range of deviation was up 100–200 mm among different products within the same category and with reference data. The results of deviation are associated with the findings of [48], who reviewed the 22 global precipitation products and was reported the magnitude of deviation up to 300 mm. All of the selected products estimated the higher amount of annual precipitation in the northern Punjab and exhibited the similar pattern from north to south gradient with the range of 1500–15 mm. However, the spatial deviation was more pronounced in the northern side as compared with the central and southern part of the study region. Moreover, the pattern of underestimation was more pronounced in the GB products as compared with SB products. One of the possible explanations of underestimation is due to lower pixel resolution of GB products over the study region [28].

Furthermore, the results showed that all of the GB products performed better in terms of lower values of ME during winter following by annual and summer monsoon season. However, the SB products showed better performance during annual scale following by summer and winter monsoon period. Overall, the agreement of GB products with observed precipitation decreases with the increase in the amount of precipitation and vice versa for SB precipitation products. The performance of SB precipitation products is well concomitant with the findings of [28,78], who reported the better performance of SB products to capture the higher magnitude of precipitation and less accurate for low precipitation events over different regions of Pakistan. In contrast, [79] revealed the better performance of SB products with less rainfall events over northwest Himalaya regions. The extent of deviation in product similarity can possibly be attributed to multiple factors, which include difference in study region, elevation, land use and land cover change [28]. The pattern similarity by using Taylor diagram indicates the better performance of GPCC and TRMM products during annual and summer monsoon seasons in terms of higher CC of around 95% during the whole study period. The results are well associated with the findings of [27], who reported the better pattern similarity performance of TRMM product in terms of higher CC of around 80% against other datasets during annual time scale over the upper Indus basin. The spatial distribution of statistical metrics (ME, RMSD, and CC) indicates the better performance of GPCC and TRMM products. However, the range of deviation in GB and SB products were more obvious in the northern Punjab with the overall pattern of underestimation (GB products) and overestimation (SB products) from the north to south gradient. The possible explanation of estimated deviation in GB and SB products over the northern Punjab could be due to wind induced errors and the presence of higher amount of aerosols in the atmosphere. The higher amount of aerosols in the atmosphere intercepts with the precipitation that reduces the efficiency of ground based meteorological stations [80]. Moreover, the wind induced errors can be attributed to the efficacy of ground based stations [81] Similarly, the pattern of overestimation in SB products particularly over northern Punjab could be associated with the higher rate of evaporation in the lower reaches of the mountains. These results are well associated with the findings of [82], who indicated the overestimation of SB products in the foothills of Himalaya mountain ranges over northern Pakistan. Moreover, the deviation in accuracy of GB products over northern Punjab could be due to inter-annual variability in weather, orographic effects and fewer gauges over the mountainous region. Many studies reported such effects in different climate studies over the northern belt of Pakistan [27,28,79]. Moreover, the fewer pixels of precipitation products also showed a weaker ability to capture the accurate spatial pattern in the central and southern parts of the study area. The possible explanation of accuracy deviation in these areas could be due the rapid extent of urban areas. Most of the meteorological station installed in the non-urban domain which could not represent the exact situation of climate variability in urban areas. Several studies documented the effect of urbanization on climate variability at a spatial scale [83].

The MK trends of observed precipitation indicate insignificant increasing trends during the period 1979–2017 and 2003–2017 with the rate of 1.12 and 5.5 mm/decade, respectively. The increasing trend was faster during period 2003–17 as compared with the period 1979–2017. The MK trend of GB and SB products showed the increasing trend with different magnitude of overestimation, except UDel and CPC products, which indicates decreasing trend as compared with observed precipitation. Furthermore, the abrupt transition analysis indicates the multiple turning points in precipitation data series during the period 1980–2005 with major abrupt changes in progressive series were detected during the period of 1984–1985 (negative), 1988–1989 (positive), 1998–1999 (negative), and 2005–2006 (positive). Many studies over the Asian and Indian monsoon system exhibited similar kind of precipitation trends over different timescales. [84–86]. The negative precipitation trend detected during the period of 1984–1985 and 1998–1999 could be due to major drought period over the whole country. The results are consistent with the findings of [87,88], who reported the severe drought conditions in the country during the mid-80s and late 90s. Moreover, the gentle increase with no obvious break point has been detected in progressive data series during the period of 2005–2017. The large inter-annual variability in reference data series were well captured by GPCC and TRMM products during the period 1979–2017 and 2003–2017, respectively.

In spite of the fact that the spatial and temporal performance of GPCC and TRMM products outperformed the respective products in terms of lower values of ME, RMSD, and higher range of CC against in-situ records during the period 1979–2017 and 2003–2017, respectively. However, it is still uncertain to estimate the accurate amount of precipitation, as witnessed by the estimated mean error in the precipitation products over the whole study region. In view of significant extent of biases in the products, we determined the factor of correction of each product for their reliable utilization in hydro-climate projects on annual, winter, and summer monsoon periods over the Punjab region. The results of estimated correction factors of different precipitation products are well covenant with the findings of [27,89], who estimated the adjustment factors of precipitation products for Upper Indus Basin (UIB). However, the magnitude of the estimated correction factors is different due to different study period and region.

The spatial and temporal performance of the products depend on a multiple factors, which are intricate in the data algorithms of the products, e.g., sources of data, spatial and temporal resolution, interpolation techniques, missing gaps in data and topography, etc. [50,90]. Similarly, the number of stations, quality, and time scale of in-situ records used for the evaluation of global products are also very important for the identification of potential product for the specific regions [72]. The detection of autocorrelation in the datasets is also important for the accuracy of trends detection and abrupt changes in the climate data series. However, the presence of autocorrelation was more significant for precipitation products, particularly for high altitude regions above 4000 m [27]. The superior performance of GPCC and TRMM precipitation products over the Punjab region might be due to their better data processing procedures, number of gauge station into account, and interpolation techniques.

The precipitation products, even with their intrinsic biases and limitations, are still important for providing valuable source of information related to precipitation variability on spatial and temporal scale. The global precipitation products are also important for climate studies when there is a lack of funding or resources to go into the field and record these types of observations. Caution must also be exercised when comparing and using the GPPs; large uncertainty exists where gauge density is low [91]. The present study highlights the potential products for precipitation over the Punjab province. The evaluation results are beneficial for improving our understanding towards the use of products in arid and semi-arid regions, like Punjab province. However, the spatial and temporal discrepancies were also identified, which will be useful for the further use of these products in hydro-meteorological applications. In view of large discrepancies among the products, future studies should focus on the new methods for the better comparative accuracy and evaluation procedure of global climate products. For instance, [92] introduce a new method of evaluation and decision making by using the fuzzy logic procedure based on the integrated linguistic operated weighted average (ILOWA) method. The main

advantage of this technique is to provide sustainable products by using different linguistic terms that are primarily an easy approach for decision making. The products assessments based on integrated models could increase the level of accuracy for the appropriate selection of potential products by considering the statistical indicators.

6. Conclusions

In this study, the gridded precipitation products (GB and SB products) were evaluated against the reference data during 1979–2017 and 2003–2017 periods on annual, winter, and summer monsoon scale by using multiple statistical methods in order to gain an understanding for the potential of Gauge and satellite based precipitation products over the Punjab province, Pakistan. The major outcomes of the study can be summarized as follows.

The result indicated that the temporal magnitude in all of the precipitation products was different and deviated up to 100–200 mm with overall spatial pattern of underestimation (GB products) and overestimation (SB products) from north to south gradient as compared with in-situ measurements. Though, the GPCC and TRMM products outperformed the relative GPPs and showed the best agreement against the reference data in terms of higher R^2 values of ≥ 0.90 and lower values of estimated RME $\leq 25\%$ at annual and summer monsoon seasons. However, CRU performed better during winter monsoon season with the estimated ME of -25.08 (-20.49%) and R-squared value of 0.91 as compared with the reference data during the period 1979–2017. All of the GB products performed better in terms of lower values of estimated ME during winter following by annual and summer monsoon season. However, SB products showed better performance during annual scale following by summer and winter monsoon period. Overall, the degree of accuracy of GB products with observed precipitation decreases with the increase in the magnitude of precipitation and vice versa for SB precipitation products

The Mann–Kendall results indicated that, the annual observed precipitation revealed the insignificant positive trend during the periods 1979–2017 and 2003–2017 with the rate of 1.12 and 5.5 mm/decade, respectively. However, the abrupt transition analysis indicates the multiple turning points in the reference data series during the period 1980–2005 with major abrupt changes were detected during the period of 1984–1985 (negative), 1988–1989 (positive), 1998–1999 (negative), and 2005–2006 (positive). The large inter-annual variability in the reference data series were well captured by GPCC and TRMM products during the period 1979–2017 and 2003–2017, respectively. In view of significant extent of biases in the GPPs, the estimated correction factor during annual, winter and summer monsoon seasons for each GPP can be useful for future hydro-climate projects over the Punjab region. The range of discrepancies in precipitation products pointed out the further improvement of GPPs for the enhancement of its accuracy over the arid and semi-arid regions, like Punjab Province. Moreover, the degree of accuracy in precipitation products varies with time and resolution; therefore, products should be assessed at different spatio-temporal scale.

References

1. Nawaz, Z.; Li, X.; Chen, Y.; Guo, Y.; Wang, X.; Nawaz, N. Temporal and spatial characteristics of precipitation and temperature in Punjab, Pakistan. *Water* **2019**, *11*, 1916. [CrossRef]
2. Jain, S.K.; Kumar, V.; Saharia, M. Analysis of rainfall and temperature trends in northeast India. *Int. J. Climatol.* **2013**, *33*, 968–978. [CrossRef]
3. Beck, H.E.; Pan, M.; Roy, T.; Weedon, G.P.; Pappenberger, F.; van Dijk, A.I.J.M.; Huffman, G.J.; Adler, R.F.; Wood, E.F. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 207–224. [CrossRef]
4. Hamdi, M.R.; Abu-Allaban, M.; Elshaieb, A.; Jaber, M.; Momani, N.M. Climate change in Jordan: A comprehensive examination approach. *Am. J. Environ. Sci.* **2009**, *5*, 740–750. [CrossRef]
5. Matti, C.; Pauling, A.; Küttel, M.; Wanner, H. Winter precipitation trends for two selected European regions over the last 500 years and their possible dynamical background. *Theor. Appl. Climatol.* **2009**, *95*, 9–26. [CrossRef]
6. Singh, P.; Kumar, V.; Thomas, T.; Arora, M. Basin-wide assessment of temperature trends in northwest and central India/Estimation par bassin versant de tendances de température au nord-ouest et au centre de l'Inde. *Hydrol. Sci. J.* **2008**, *53*, 421–433. [CrossRef]
7. De la Casa, A.; Nasello, O. Breakpoints in annual rainfall trends in Córdoba, Argentina. *Atmos. Res.* **2010**, *95*, 419–427. [CrossRef]
8. Gleick, P.H. Climate change, hydrology, and water resources. *Rev. Geophys.* **1989**, *27*, 329–344. [CrossRef]
9. Feng, S.; Hu, Q.; Qian, W. Quality control of daily meteorological data in China, 1951–2000: A new dataset. *Int. J. Climatol. A J. R. Meteorol. Soc.* **2004**, *24*, 853–870. [CrossRef]
10. Jiang, S.; Ren, L.; Hong, Y.; Yong, B.; Yang, X.; Yuan, F.; Ma, M. Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *J. Hydrol.* **2012**, *452*, 213–225. [CrossRef]
11. Liu, X.; Yang, T.; Hsu, K.; Liu, C.; Sorooshian, S. Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 169–181. [CrossRef]
12. Tapiador, F.J.; Turk, F.J.; Petersen, W.; Hou, A.Y.; García-Ortega, E.; Machado, L.A.T.; Angelis, C.F.; Salio, P.; Kidd, C.; Huffman, G.J. Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.* **2012**, *104*, 70–97. [CrossRef]
13. Wang, W.; Lu, H.; Zhao, T.; Jiang, L.; Shi, J. Evaluation and comparison of daily rainfall from latest GPM and TRMM products over the Mekong River Basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2540–2549. [CrossRef]
14. Nashwan, M.; Shahid, S.; Chung, E.-S.; Ahmed, K.; Song, Y. Development of climate-based index for hydrologic hazard susceptibility. *Sustainability* **2018**, *10*, 2182. [CrossRef]
15. Albright, T.P.; Pidgeon, A.M.; Rittenhouse, C.D.; Clayton, M.K.; Wardlow, B.D.; Flather, C.H.; Culbert, P.D.; Radeloff, V.C. Combined effects of heat waves and droughts on avian communities across the conterminous United States. *Ecosphere* **2010**, *1*, 1–22. [CrossRef]
16. Nkiaka, E.; Nawaz, N.R.; Lovett, J.C. Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: Case study of the Logone catchment, Lake Chad Basin. *Meteorol. Appl.* **2017**, *24*, 9–18. [CrossRef]
17. Zhao, Y.; Xie, Q.; Lu, Y.; Hu, B. Hydrologic evaluation of TRMM multisatellite precipitation analysis for Nanliu river basin in Humid Southwestern China. *Sci. Rep.* **2017**, *7*, 2470. [CrossRef]
18. Kunkel, K.E.; Vose, R.S.; Stevens, L.E.; Knight, R.W. Is the monthly temperature climate of the United States becoming more extreme? *Geophys. Res. Lett.* **2015**, *42*, 629–636. [CrossRef]
19. Kunkel, K.E.; Andsager, K.; Easterling, D.R. Long-term trends in extreme precipitation events over the conterminous United States and Canada. *J. Clim.* **1999**, *12*, 2515–2527. [CrossRef]
20. Rajah, K.; O'Leary, T.; Turner, A.; Petrakis, G.; Leonard, M.; Westra, S. Changes to the temporal distribution of daily precipitation. *Geophys. Res. Lett.* **2014**, *41*, 8887–8894. [CrossRef]
21. Daly, C.; Slater, M.E.; Roberti, J.A.; Laseter, S.H.; Swift, L.W., Jr. High-resolution precipitation mapping in a mountainous watershed: Ground truth for evaluating uncertainty in a national precipitation dataset. *Int. J. Climatol.* **2017**, *37*, 124–137. [CrossRef]

22. Hunziker, S.; Gubler, S.; Calle, J.; Moreno, I.; Andrade, M.; Velarde, F.; Ticona, L.; Carrasco, G.; Castellón, Y.; Oria, C. Identifying, attributing, and overcoming common data quality issues of manned station observations. *Int. J. Climatol.* **2017**, *37*, 4131–4145. [CrossRef]

23. Khairul, I.; Mastrantonas, N.; Rasmy, M.; Koike, T.; Takeuchi, K. Inter-comparison of gauge-corrected global satellite rainfall estimates and their applicability for effective water resource management in a transboundary river basin: The case of the Meghna River basin. *Remote Sens.* **2018**, *10*, 828. [CrossRef]

24. Samadi, S.; Carbone, G.J.; Mahdavi, M.; Sharifi, F.; Bihamta, M.R. Statistical downscaling of climate data to estimate streamflow in a semi-arid catchment. *Hydrol. Earth Syst. Sci. Discuss.* **2012**, *9*, 4869–4918. [CrossRef]

25. Buytaert, W.; Friesen, J.; Liebe, J.; Ludwig, R. Assessment and management of water resources in developing, semi-arid and arid regions. *Water Resour. Manag.* **2012**, *26*, 841–844. [CrossRef]

26. Ahmad, W.; Fatima, A.; Awan, U.K.; Anwar, A. Analysis of long term meteorological trends in the middle and lower Indus Basin of Pakistan—A non-parametric statistical approach. *Glob. Planet. Chang.* **2014**, *122*, 282–291. [CrossRef]

27. Dahri, Z.H.; Ludwig, F.; Moors, E.; Ahmad, B.; Khan, A.; Kabat, P. An appraisal of precipitation distribution in the high-altitude catchments of the Indus basin. *Sci. Total Environ.* **2016**, *548*, 289–306. [CrossRef]

28. Ullah, W.; Wang, G.; Ali, G.; Tawia Hagan, D.F.; Bhatti, A.S.; Lou, D. Comparing multiple precipitation products against in-situ observations over different climate regions of Pakistan. *Remote Sens.* **2019**, *11*, 628. [CrossRef]

29. Krakauer, N.Y.; Lakhankar, T.; Dars, G.H. Precipitation trends over the Indus basin. *Climate* **2019**, *7*, 116. [CrossRef]

30. Iqbal, M.F.; Athar, H. Validation of satellite based precipitation over diverse topography of Pakistan. *Atmos. Res.* **2018**, *201*, 247–260. [CrossRef]

31. Khattak, M.S.; Reman, N.U.; Sharif, M.; Khan, M.A. Analysis of streamflow data for trend detection on major rivers of the Indus Basin. *J. Himal. Earth Sci.* **2015**, *48*, 87.

32. Ashiq, M.W.; Zhao, C.; Ni, J.; Akhtar, M. GIS-based high-resolution spatial interpolation of precipitation in mountain–plain areas of Upper Pakistan for regional climate change impact studies. *Theor. Appl. Climatol.* **2010**, *99*, 239. [CrossRef]

33. Abbas, F. Analysis of a historical (1981–2010) temperature record of the Punjab province of Pakistan. *Earth Interact.* **2013**, *17*, 1–23. [CrossRef]

34. Khan, S.U.; Hasan, M.U.; Khan, F.K.; Bari, A. Climate classification of Pakistan. In Proceedings of the Fourth International Scientific Conference BALWOIS 2010, Ohrid, North Macedonia, 25–29 May 2010.

35. Asmat, U.; Athar, H.; Nabeel, A.; Latif, M. An AOGCM based assessment of interseasonal variability in Pakistan. *Clim. Dyn.* **2018**, *50*, 349–373. [CrossRef]

36. Nawaz, Z.; Li, X.; Chen, Y.; Wang, X.; Zhang, K.; Nawaz, N.; Guo, Y.; Meerzhan, A. Spatiotemporal assessment of temperature data products for the detection of warming trends and abrupt transitions over the largest irrigated area of Pakistan. *Adv. Meteorol.* **2020**, *2020*, 3584030. [CrossRef]

37. Chen, Y.; Ebert, E.E.; Walsh, K.J.E.; Davidson, N.E. Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN data. *J. Geophys. Res. Atmos.* **2013**, *118*, 2184–2196. [CrossRef]

38. 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]

39. Iqbal, M.; Wen, J.; Wang, X.; Lan, Y.; Tian, H.; Anjum, M.N.; Adnan, M. Assessment of air temperature trends in the source Region of Yellow River and its sub-basins. *Asia Pac. J. Atmos. Sci.* **2018**, *54*, 111–123. [CrossRef]

40. Mitchell, J.M.; Dzerdzeevskii, B.; Flohn, H.; Hofmeyr, W.L.; Lamb, H.H.; Rao, K.N.; Wallén, C.C. *Climatic Change*; Technical Note, No. 79; World Meteorological Organization: Geneva, Switzerland, 1966; Volume 99.

41. Peterson, T.C.; Easterling, D.R.; Karl, T.R.; Groisman, P.; Nicholls, N.; Plummer, N.; Torok, S.; Auer, I.; Boehm, R.; Gullett, D. Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol. A J. R. Meteorol. Soc.* **1998**, *18*, 1493–1517. [CrossRef]

42. Kohler, M.A. On the Use of Double-Mass Analysis for Testing the Consistency of Meteorological Records and for Making Required Adjustments. *Bull. Am. Meteorol. Soc. J.* **1949**, *30*, 188–195. [CrossRef]

43. Craven, P.; Wahba, G. Smoothing noisy data with spline functions. *Numer. Math.* **1978**, *31*, 377–403. [CrossRef]

44. Hutchinson, M.F.; Gessler, P.E. Splines—More than just a smooth interpolator. *Geoderma* **1994**, *62*, 45–67. [CrossRef]

45. Partal, T. Wavelet transform-based analysis of periodicities and trends of Sakarya basin (Turkey) streamflow data. *River Res. Appl.* **2010**, *26*, 695–711. [CrossRef]

46. Sonali, P.; Kumar, D.N. Review of trend detection methods and their application to detect temperature changes in India. *J. Hydrol.* **2013**, *476*, 212–227. [CrossRef]

47. Sayemuzzaman, M.; Jha, M.K. Seasonal and annual precipitation time series trend analysis in North Carolina, United States. *Atmos. Res.* **2014**, *137*, 183–194. [CrossRef]

48. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.* **2018**, *56*, 79–107. [CrossRef]

49. Willmott, C.J.; Matsuura, K. Smart interpolation of annually averaged air temperature in the United States. *J. Appl. Meteorol.* **1995**, *34*, 2577–2586. [CrossRef]

50. Harris, I.; Jones, P.D.; Osborn, T.J.; Lister, D.H. Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *Int. J. Climatol.* **2014**, *34*, 623–642. [CrossRef]

51. Yatagai, A.; Kamiguchi, K.; Arakawa, O.; Hamada, A.; Yasutomi, N.; Kitoh, A. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1401–1415. [CrossRef]

52. Xie, P.; Chen, M.; Shi, W. CPC unified gauge-based analysis of global daily precipitation. In Proceedings of the 24th Conference on Hydrology, Atlanta, GA, USA, 16–21 January 2010.

53. 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]

54. Li, W.; Sun, W.; He, X.; Scaioni, M.; Yao, D.; Chen, Y.; Gao, J.; Li, X.; Cheng, G. Improving CHIRPS daily satellite-precipitation products using coarser ground observations. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1678–1682. [CrossRef]

55. Ashouri, H.; Hsu, K.-L.; Sorooshian, S.; Braithwaite, D.K.; Knapp, K.R.; Cecil, L.D.; Nelson, B.R.; Prat, O.P. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 69–83. [CrossRef]

56. Adler, R.F.; Huffman, G.J.; Chang, A.; Ferraro, R.; Xie, P.-P.; Janowiak, J.; Rudolf, B.; Schneider, U.; Curtis, S.; Bolvin, D. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **2003**, *4*, 1147–1167. [CrossRef]

57. Hong, Y.; Gochis, D.; Cheng, J.; Hsu, K.; Sorooshian, S. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *J. Hydrometeorol.* **2007**, *8*, 469–482. [CrossRef]

58. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 1–21. [CrossRef]

59. Ushio, T.; Sasashige, K.; Kubota, T.; Shige, S.; Okamoto, K.; Aonashi, K.; Inoue, T.; Takahashi, N.; Iguchi, T.; Kachi, M. A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteorol. Soc. Japan. Ser. II* **2009**, *87*, 137–151. [CrossRef]

60. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]

61. Chen, Y.; Guan, Y.; Shao, G.; Zhang, D. Investigating trends in streamflow and precipitation in Huangfuchuan Basin with wavelet analysis and the Mann-Kendall test. *Water* **2016**, *8*, 77. [CrossRef]

62. Yang, P.; Xia, J.; Zhang, Y.; Hong, S. Temporal and spatial variations of precipitation in Northwest China during 1960–2013. *Atmos. Res.* **2017**, *183*, 283–295. [CrossRef]

63. Zamani, R.; Mirabbasi, R.; Abdollahi, S.; Jhajharia, D. Streamflow trend analysis by considering autocorrelation structure, long-term persistence, and Hurst coefficient in a semi-arid region of Iran. *Theor. Appl. Climatol.* **2017**, *129*, 33–45. [CrossRef]

64. Sen, P.K. Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* **1968**, *63*, 1379–1389. [CrossRef]

65. Zhang, X.; Zwiers, F.W.; Li, G. Monte Carlo experiments on the detection of trends in extreme values. *J. Clim.* **2004**, *17*, 1945–1952. [CrossRef]

66. Sneyers, R. *On the Statistical Analysis of Series of Observations*; World Meteorological Organization: Geneva, Switzerland, 1990; ISBN 9263104158.

67. Zarenistanak, M.; Dhorde, A.G.; Kripalani, R.H. Trend analysis and change point detection of annual and seasonal precipitation and temperature series over southwest Iran. *J. Earth Syst. Sci.* **2014**, *123*, 281–295. [CrossRef]

68. Mann, H.B. Non-parametric tests against trend. *Econmetrica* **1945**, *13*, 245–259. [CrossRef]

69. Mantua, N.J.; Hare, S.R.; Zhang, Y.; Wallace, J.M.; Francis, R.C. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.* **1997**, *78*, 1069–1080. [CrossRef]

70. Kendall, M.G. *Rank Correlation Measures*; Charles Griffin: London, UK, 1975; Volume 202, p. 15.

71. Ullah, S.; You, Q.; Ullah, W.; Ali, A. Observed changes in precipitation in China-Pakistan economic corridor during 1980–2016. *Atmos. Res.* **2018**, *210*, 1–14. [CrossRef]

72. Ahmed, K.; Shahid, S.; Wang, X.; Nawaz, N.; Khan, N. Evaluation of gridded precipitation datasets over arid regions of Pakistan. *Water* **2019**, *11*, 210. [CrossRef]

73. 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. Sp. Res.* **2014**, *54*, 670–684. [CrossRef]

74. Xu, M.; Kang, S.; Wu, H.; Yuan, X. Detection of spatio-temporal variability of air temperature and precipitation based on long-term meteorological station observations over Tianshan Mountains, Central Asia. *Atmos. Res.* **2018**, *203*, 141–163. [CrossRef]

75. Rashid, M.M.; Beecham, S.; Chowdhury, R.K. Assessment of trends in point rainfall using continuous wavelet transforms. *Adv. Water Resour.* **2015**, *82*, 1–15. [CrossRef]

76. Priya, P.; Mujumdar, M.; Sabin, T.P.; Terray, P.; Krishnan, R. Impacts of Indo-Pacific sea surface temperature anomalies on the summer monsoon circulation and heavy precipitation over northwest India–Pakistan region during 2010. *J. Clim.* **2015**, *28*, 3714–3730. [CrossRef]

77. Gampe, D.; Ludwig, R. Evaluation of gridded precipitation data products for hydrological applications in complex topography. *Hydrology* **2017**, *4*, 53. [CrossRef]

78. Naveed, M.; Ding, Y.; Shangguan, D.; Ahmad, I. Performance evaluation of latest integrated multi-satellite retrievals for Global Precipitation Measurement (IMERG) over the northern highlands of Pakistan. *Atmos. Res.* **2018**, *205*, 134–146.

79. Bharti, V.; Singh, C. Evaluation of error in TRMM 3B42V7 precipitation estimates over the Himalayan region. *J. Geophys. Res. Atmos.* **2015**, *120*, 12458–12473. [CrossRef]

80. Kaufman, Y.J.; Tanré, D.; Boucher, O. A satellite view of aerosols in the climate system. *Nature* **2002**, *419*, 215–223. [CrossRef] [PubMed]

81. Tahir, A.A.; Adamowski, J.F.; Chevallier, P.; Haq, A.U.; Terzago, S. Comparative assessment of spatiotemporal snow cover changes and hydrological behavior of the Gilgit, Astore and Hunza River basins (Hindukush–Karakoram–Himalaya region, Pakistan). *Meteorol. Atmos. Phys.* **2016**, *128*, 793–811. [CrossRef]

82. Palazzi, E.; von Hardenberg, J.; Provenzale, A. Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *J. Geophys. Res. Atmos.* **2013**, *118*, 85–100. [CrossRef]

83. Ren, G.; Zhou, Y. Urbanization effect on trends of extreme temperature indices of national stations over Mainland China, 1961–2008. *J. Clim.* **2014**, *27*, 2340–2360. [CrossRef]

84. Kripalani, R.H.; Oh, J.H.; Kulkarni, A.; Sabade, S.S.; Chaudhari, H.S. South Asian summer monsoon precipitation variability: Coupled climate model simulations and projections under IPCC AR4. *Theor. Appl. Climatol.* **2007**, *90*, 133–159. [CrossRef]

85. Latif, M.; Syed, F.S.; Hannachi, A. Rainfall trends in the South Asian summer monsoon and its related large-scale dynamics with focus over Pakistan. *Clim. Dyn.* **2017**, *48*, 3565–3581. [CrossRef]

86. Wang, P.X.; Wang, B.; Cheng, H.; Fasullo, J.; Guo, Z.; Kiefer, T.; Liu, Z. The global monsoon across time scales: Mechanisms and outstanding issues. *Earth Sci. Rev.* **2017**, *174*, 84–121. [CrossRef]

87. Xie, H.; Ringler, C.; Zhu, T.; Waqas, A. Droughts in Pakistan: A spatiotemporal variability analysis using the Standardized Precipitation Index. *Water Int.* **2013**, *38*, 620–631. [CrossRef]

88. Jamro, S.; Dars, G.H.; Ansari, K.; Krakauer, N.Y. Spatio-temporal variability of drought in Pakistan using standardized precipitation evapotranspiration index. *Appl. Sci.* **2019**, *9*, 4588. [CrossRef]

89. Lutz, A.F.; Immerzeel, W.W.; Shrestha, A.B.; Bierkens, M.F.P. Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nat. Clim. Chang.* **2014**, *4*, 587–592. [CrossRef]

90. Fu, G.; Yu, J.; Yu, X.; Ouyang, R.; Zhang, Y.; Wang, P.; Liu, W.; Min, L. Temporal variation of extreme rainfall events in China, 1961–2009. *J. Hydrol.* **2013**, *487*, 48–59. [CrossRef]

91. McEvoy, D.J.; Mejia, J.F.; Huntington, J.L. Use of an observation network in the Great Basin to evaluate gridded climate data. *J. Hydrometeorol.* **2014**, *15*, 1913–1931. [CrossRef]

92. Rodger, J.A.; George, J.A. Triple bottom line accounting for optimizing natural gas sustainability: A statistical linear programming fuzzy ILOWA optimized sustainment model approach to reducing supply chain global cybersecurity vulnerability through information and communications t. *J. Clean. Prod.* **2017**, *142*, 1931–1949. [CrossRef]

Article

Weather Types Affect Rain Microstructure: Implications for Estimating Rain Rate

Wael Ghada [1,*], Joan Bech [2], Nicole Estrella [1], Andreas Hamann [3] and Annette Menzel [1,4]

[1] TUM School of Life Sciences, Technical University of Munich, Hans-Carl-von-Carlowitz-Platz 2,
    D-85354 Freising, Germany; estrella@wzw.tum.de (N.E.); annette.menzel@tum.de (A.M.)

[2] Department of Applied Physics-Meteorology, University of Barcelona, C/ Marti i Franques 1,
    08028 Barcelona, Spain; joan.bech@ub.edu

[3] Department of Renewable Resources, University of Alberta, 733 General Services Building,
    Edmonton, AB T6G 2H1, Canada; andreas.hamann@ualberta.ca

[4] Institute for Advanced Study, Technical University of Munich, Lichtenbergstraße 2a,
    D-85748 Garching, Germany

* Correspondence: ghada@wzw.tum.de; Tel.: +49-81-6171-4743

Abstract: Quantitative precipitation estimation (QPE) through remote sensing has to take rain microstructure into consideration, because it influences the relationship between radar reflectivity Z and rain intensity R. For this reason, separate equations are used to estimate rain intensity of convective and stratiform rain types. Here, we investigate whether incorporating synoptic scale meteorology could yield further QPE improvements. Depending on large-scale weather types, variability in cloud condensation nuclei and the humidity content may lead to variation in rain microstructure. In a case study for Bavaria, we measured rain microstructure at ten locations with laser-based disdrometers, covering a combined 18,600 h of rain in a period of 36 months. Rain was classified on a temporal scale of one minute into convective and stratiform based on a machine learning model. Large-scale wind direction classes were on a daily scale to represent the synoptic weather types. Significant variations in rain microstructure parameters were evident not only for rain types, but also for wind direction classes. The main contrast was observed between westerly and easterly circulations, with the latter characterized by smaller average size of drops and a higher average concentration. This led to substantial variation in the parameters of the radar rain intensity retrieval equation Z–R. The effect of wind direction on Z–R parameters was more pronounced for stratiform than convective rain types. We conclude that building separate Z–R retrieval equations for regional wind direction classes should improve radar-based QPE, especially for stratiform rain events.

Keywords: Thies; disdrometer; weather circulations; convective; stratiform; rain spectra; radar reflectivity–rain rate relationship

---

1. Introduction

Understanding rain microstructure can provide us with an insight into the prevailing rain formation processes leading to it. This understanding can be employed in improving quantitative estimation of rain intensity using weather radar, especially in flat regions with high altitude values of the zero degree isotherm [1–4]. Furthermore, the parametrization of the microphysical processes in numerical weather and climate models can be improved [5,6]. Rain microstructure varies on different spatial scales ranging from few meters [7], to few hundreds of meters [8], to regional [9,10] and global extents [11,12]. This variation also occurs with seasons [13], rain types [14], and large-scale weather types [15–17].

A clear example of the different rain formation processes leading to variations in rain drop size distribution is the discrepancy between convective and stratiform rain. This has been quantified in a number of studies [5,14,18,19]. The reason for the difference is the relative importance of cold and warm rain formation processes [20]. Stratiform rain is formed mainly by processes involving ice crystals and interactions of ice with liquid water, while convective rain formation comprises both warm and cold processes. Factors and processes that influence the rain drop size distribution as observed on the ground include rimming and aggregation (above the 0 °C isotherm), condensation (below the 0 °C isotherm), collision, coalescence, turbulence, cloud thickness, electric field, evaporation, and drop fragmentation [21,22]. The difference in rain drop size distribution between convective rain and stratiform rain has been used for the classification of both rain types on the ground. Most of these methods use two rain drop size distribution parameters and a linear discrimination between the regions of rain types [19,23–26]. Recent methods employed machine learning and reached higher performance levels when using four rain drop size distribution parameters [27,28].

Large-scale weather types denote atmospheric conditions such as the high and low pressure distribution, the position and paths of frontal zones, and the existence of cyclonic or anticyclonic circulation types over a sequence of days [29]. Indirectly, they also influence stream flows [30], floods [31–33], debris-flow events [34], forest fires [35,36], air quality, and pollen distribution [37–39].

Weather type classification is an important part of statistical climatology [40,41], because these types explain many local weather phenomena. Weather types influence local near-surface temperatures and precipitation [42–46]. They also affect the diurnal cycle of precipitation in terms of frequency and amount [47–49], and they impact the occurrence and the magnitude of meteorological extreme events [50–54]. Large-scale weather types may therefore also influence rain microstructure by different rain formation processes being more prevalent under different synoptic scale conditions.

Quantifying rain microstructure under different large-scale weather types may have practical applications for radar-based estimation of rain intensity, because the microstructure influences the relationship between radar reflectivity Z and rain intensity R. For this reason, separate equations are used to estimate rain intensity of convective and stratiform rain type [10,55], instead of using one equation that fits both rain types. A similar improvement of the radar estimation of rain might be possible when considering specific Z–R relations for each of the weather types. We previously reported weather type specific Z–R models with lower errors in estimating rain intensity in Lausanne, Switzerland [17]. Similarly, the influence of weather types on Z–R relationships was also reported for the Cévennes-Vivarais Region, France [16]. However, parameterizing Z–R equations for many weather types definitively requires large amounts of data to represent each class.

Here, we contribute an analysis of the relationship between Z–R parameters and weather types in Central Europe, based on a comprehensive regional dataset of rain microstructure measurements at ten sites in the federal state of Bavaria, Germany. We ask: (1) What is the effect of weather types on rain microstructure, considering both types of rain? and (2) Is there consistent variation in the Z–R parameters between weather types that would suggest opportunities to improve QPE with radar-based methods? To address these questions, we investigate disdrometer records under different large-scale wind direction patterns at a daily scale, and rain type classifications at one-minute intervals over a period of three years.

2. Materials and Methods

2.1. Data Sources and Tools

We obtained raw rain drop size distribution measurements from the German Meteorological Service (Deutscher Wetterdienst, DWD), operating a network of Thies disdrometers in Bavaria, in the southeast of Germany (Figure 1). We analyzed measurements at ten sites spanning a period of three years (January 2014–December 2016) with a temporal resolution of one minute. The disdrometers locations cover a distance of 167 km from north to south and 185 km from east to west.

Figure 1. Disdrometer locations in Bavaria (SE Germany) that were used to measure rain microstructure, covering a total of 18,600 h of rain in a period of 36 months.

Since raw disdrometer data requires some statistical data cleaning procedures to remove erroneous readings, we followed the filtering procedure of Friedrich et al. [56] and the additional steps of Ghada et al. [17] to remove unrealistically large particles, margin fallers, splashing effects, or readings of insect and spider webs. The filtering procedure removed: (1) All measured particles with a diameter larger than 8 mm; (2) All particles which had a falling velocity less than 60% or greater than 140% of the terminal velocity associated with rain drops of the corresponding diameter [57,58] (Figure 2); (3) Intervals marked by a damaged laser signal or as non-rain intervals by the disdrometer; (4) Intervals which included large drops (D > 5 mm) with low velocities (V < 1 m/h) as an indicator of high wind speed; (5) Intervals with rain intensity lower than 0.1 mm/h [59,60]; (6) Intervals with three or less diameter bins to insure the existence of a drop size distribution. After filtering, the dataset contained a total of 21,705 mm of accumulated rain over a period of 18,633 h.
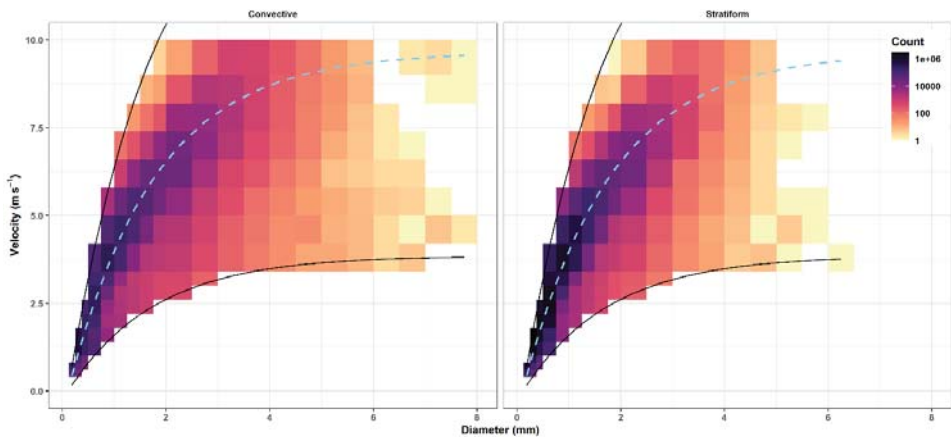
Figure 2. Raindrop count in each diameter-velocity range after the filtering process. The dotted line represents the terminal velocity of each diameter value. The solid lines represent the 60% and the 140% of the terminal velocity.

The DWD classifies large-scale synoptic weather patterns into 40 classes of weather types. The weather type is provided on a daily time scale and is applicable to all of Germany and its

surroundings. The classification is based on an operational numerical weather prediction system, i.e., modelling different atmospheric fields such as geopotential height, temperature, relative humidity, and the zonal and meridional components of the wind for several elevations. A detailed explanation of the classification procedure is available online [61], and the full record of weather types is provided by the DWD [62]. Since this classification is performed on daily basis, it would be operationally feasible to associate a separate configuration of the radar rain rate estimate for each weather type class. However, in order to simplify the classification for the purpose of this exploratory case, we grouped all possible classes according to their wind direction index. This index takes one of five possible values: northeasterly (NE), southeasterly (SE), southwesterly (SW), northwesterly (NW), and no prevailing direction (XX). Determining the specific wind direction is based on the number of grid points over Germany with a specific wind direction which needs to exceed 2/3 of the total number of grid points. In case this threshold was not exceeded, the wind direction index is assigned to XX.

For data filtering, analysis, and production of visual and statistical results, we used R [63], RStudio [64], and the packages caret [65], e1071 [66], reshape2 [67], raster [68], Rmisc [69], ggplot2 [70], and rnaturalearth [71].

2.2. Drop Size Distribution Parameters

Thies disdrometers are laser-based instruments that provide high temporal records of rain microstructure. When a precipitation particle passes between the transmitter and the receiver, the strength of the laser beam is reduced. Based on the magnitude and duration of this reduction, it is possible to estimate the size and velocity of the passing precipitation particle. The Thies disdrometers raw data output represents one-minute summaries of the number of particles in 22 non-linear size classes and 20 non-linear velocity classes. From the raw output, a number of parameters can be obtained. This study is focused particularly on rain intensity R, radar reflectivity Z, total number of drop concentration N, and median volume drop diameter D0.

Rain rate R (mm/h) is given by

$$R = \frac{6 \times 10^{-4} \times \pi}{\Delta T} \times \sum_{i=1}^{i=22} \sum_{j=1}^{j=20} \left(x_{i,j} \frac{D_i^3}{A_i} \right) \tag{1}$$

where

$x_{i,j}$: Detected number of drops that fall in diameter range i and velocity range j,

ΔT (s): Temporal resolution (60 s in this case),

A_i (m²): Corrected detection area: $A_i = 228 \times \left(20 - \frac{D_i}{2} \right)/10^6$,

D_i (mm): Mean diameter of drops that fall in diameter range i.

The radar reflectivity Z (dBZ) is calculated with the following expression:

$$Z = 10 * log_{10} \left(\sum_{i=1}^{i=22} \sum_{j=1}^{j=20} \left(x_{i,j} \frac{D_i^6}{\left(A_i \, V_j \, \Delta T \right)} \right) \right) \tag{2}$$

where V_j (m/s) : Mean velocity of drops that fall in the velocity range j.

The total number of drops N (m^{-3}) is computed according to

$$N = \sum_{i=1}^{i=22} \sum_{j=1}^{j=20} \left(\frac{x_{i,j}}{A_i \, V_j \, W_i \, \Delta T} \right) \tag{3}$$

where W_i (mm): the width of the diameter range i.

The rain microstructure is assumed to follow a gamma distribution [72]:

$$N(D) = N_0 D^\mu e^{(-\Lambda D)} \tag{4}$$

where $N(D)$ (mm^{-1}m^{-3}) is the number of drops for each diameter range per unit volume and unite size. The intercept N_0 (mm$^{-1-\mu}$ m^{-3}), the shape μ (-), and the slope Λ (mm^{-1}) parameters are determined by the moments method [73]. The nth moment of the raindrop size distribution M_n (mm$^{-1-\mu}$ m^{-3}) is given by

$$M_n = \int_{D_{min}}^{D_{max}} D^n N(D) dD \tag{5}$$

and the three gamma parameters are

$$N_0 = \frac{\Lambda^{\mu+3} M_2}{\Gamma(\mu+3)} \tag{6}$$

$$\mu = \frac{(7-11\eta) - \left[(7-11\eta)^2 - 4(\eta-1)(30\eta-12)\right]^{0.5}}{2(\eta-1)} \tag{7}$$

$$\Lambda = \left[\frac{(4+\mu)(3+\mu)M_2}{M_4}\right]^{0.5} \tag{8}$$

where

$$\eta = \frac{M_4{}^2}{M_2 M_6} \tag{9}$$

The mass weighted mean diameter D_m (mm), the median volume diameter D_0 (mm) and the normalized intercept N_w (mm^{-1}m^{-3}) are calculated based on the parameters of gamma distribution:

$$D_m = \frac{M_4}{M_3} \tag{10}$$

$$D_0 = \frac{D_m(\mu + 3.67)}{\mu + 4} \tag{11}$$

$$N_w = \frac{4^4 \, M_3}{6 \, D_m{}^4} \tag{12}$$

Additionally, the classification of rain type into convective and stratiform requires the use of the following parameters: sd_N_10, sd_D$_0$_10, and sd_log$_{10}$_R_10, where sd_XX_10 is the standard deviation of the values of XX (XX being N, D$_0$ and R, respectively) over a time window of ten minutes.

2.3. Rain Type Classification

Rain type classification was based on an ensemble classifier to predict stratiform versus convective rain based on cloud type, rain intensity, and the standard deviation of rain intensities calculated over the span of ten minutes.

To create a training set for the machine learning model that classifies rain type into convective and stratiform, we obtained records of cloud genera from the DWD [74]. These ground observations were available between July 2013 and August 2014 at Fürstenzell and between July 2013 and January 2014 at Regensburg.

A random forest classification model was trained on the available data from the two locations in this dataset. A combination of two criteria was used for the prior classification, the observation of cloud genus, and the values of R and its standard deviation over five minutes. The model was trained based on the intervals where the prior classification was feasible. It was then used to classify rain in the whole dataset. The spatial variability in rain properties might influence the quality of our classification scheme, especially that the model was trained in only two out of the ten sites. However, the drop in

quality on this scale when training in one location and testing in another was minor [28]. More details about the classification procedure are given in Ghada et al [28].

2.4. Retrieving the Parameters of the Z–R Relation

Weather radars usually provide the reflectivity Z which is transformed into rain intensity R using an exponential equation. In our case, R and Z are provided by the disdrometer; therefore, it is possible to get the values of A and b by fitting a linear model to the values of $log_{10}(R)$ and Z.

The radar reflectivity Z is assumed to be related to rain intensity R by the power law:

$$Z = A \times R^b \tag{13}$$

In this equation, Z is expressed in $mm^6\ m^{-3}$. However, Z is usually expressed in the unit decibel relative to Z (dBZ):

$$Z_{[dBZ]} = 10 \times log_{10}\left(Z_{[mm^6\ m^{-3}]}\right). \tag{14}$$

By taking the log of Equation (13) and multiplying by 10:

$$10 \times log_{10}(Z) = 10 \times log_{10}(A) + 10 \times b \times log_{10}(R) \tag{15}$$

Moreover, based on Equation (14):

$$dBZ = 10 \times log_{10}(A) + 10 \times b \times log_{10}(R) \tag{16}$$

a simple linear model is fitted to the values of dBZ and log R which are calculated from the rain drop size distribution. This linear model has the equation:

$$dBZ = intercept + slope \times log_{10}(R) \tag{17}$$

thus, by comparing Equations (16) and (17) the A and b parameters can be readily found:

$$b = \frac{slope}{10} \tag{18}$$

$$A = 10^{\frac{intercept}{10}} \tag{19}$$

Equations (13)–(19) represent the conventional way of retrieving A and b. An alternative method is to consider R as the dependent variable [75]. This method is more appropriate because the main purpose is to reduce errors in estimating R:

$$R = (1/A)^{1/b} \times Z^{1/b} \tag{20}$$

By taking the log_{10} of both sides of Equation (20):

$$log_{10}(R) = \frac{1}{b} \times log_{10}(Z) - \frac{1}{b} \times log_{10}(A) \tag{21}$$

$$log_{10}(R) = \frac{dBZ}{10 \times b} - \frac{log_{10}(A)}{b} \tag{22}$$

$$log_{10}(R) = intercept + slope \times dBZ \tag{23}$$

by comparing Equations (22) and (23):

$$b = \frac{1}{slope \times 10} \tag{24}$$

$$A = 10^{-b \times intercept} \tag{25}$$

Retrieval of A and b values was done for each event separately. Events with an accumulated rain amount of less than 1 mm were excluded to limit their influence on the fitting process. Additionally, the events were defined by a minimum interevent threshold of 15 min and a minimum duration of 15 min as in Jaffrain and Berne [75]. To ensure clear classification of the rain type on the event level, the fitting was restricted to events during which more than 60% of the event was convective, and events where all intervals were classified as stratiform. The remaining 2449 events contain 9914 h of rain (see Table A1).

3. Results

3.1. Duration and Amount Variation With Rain Type and Wind Direction

During the 1096 days included in the study period, rain was recorded at least at one station on 515 days. The five wind directions had different frequencies and the most frequent wind directions were the westerly circulations SW and NW with a total of 739 days or two thirds of the time (Figure 3). More than half of these days included rain in at least one station. The easterly circulations accounted for less than 12% of the total number of days. SE had the lowest occurrence and the lowest percentage of rainy days. Both XX and NE had more than 40% rainy days.
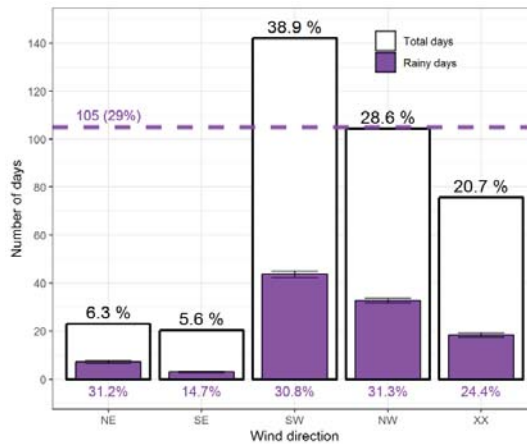
Figure 3. Frequency of rainy days per year and per wind direction classes that represent large-scale weather types. Rainy days are days on which at least one station recorded five minutes of rain with an intensity of more than 0.2 mm/h. Error bars represent the 95% confidence intervals. Percentages above the white columns represent the overall occurrence of each wind direction and percentages below the columns represent the portion of rainy days in the total number of days within a specific wind direction. The dashed line represents the mean number of rainy days per year.

When examining the accumulated rain amount and duration, westerly circulations were the dominant wind directions with a contribution reaching 69% of the total rain duration (18,633 h) and total rain amount (21,705 mm) accumulated over all stations (Figure 4). Easterly circulations contributed less than 10% of both rain duration and amount. Convection contributed 36% of the total rain amount and occupied only 8.5% of rain duration. Southerly circulations had the highest proportion of convective rain with around 10% of the total rain duration and more than 40% of the total rain amount, while northerly, and especially northeasterly circulations had a low proportion of convective rain.
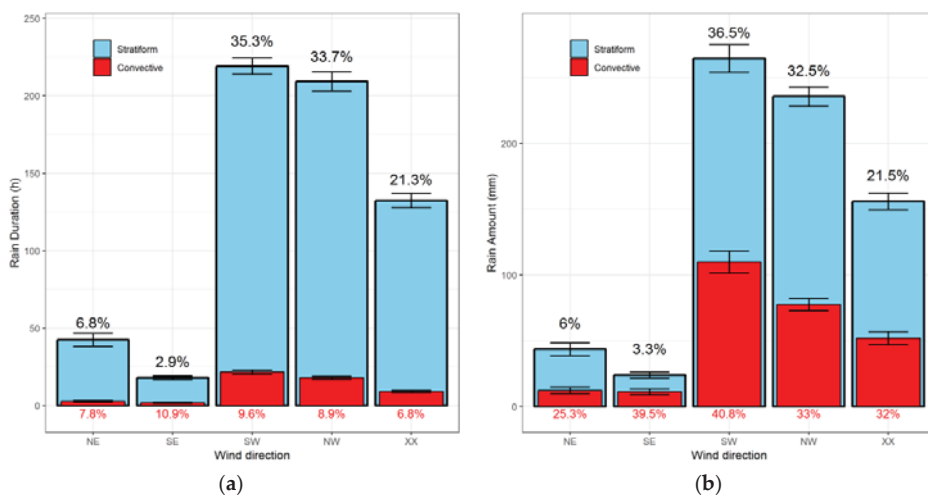
Figure 4. Prevalence of convective and stratiform rain types. Accumulated rain duration (**a**), and rain amount (**b**) per wind direction averaged over the stations and years. Error bars represent the 95% confidence intervals. The percentages on top of each column represent the proportion of accumulated rain within the respective wind direction to the accumulated rain in the whole year. The percentages below the columns represent the proportion of convective rain to total rain within the respective wind direction.

The mean stratiform rain intensity was 0.8 mm/h which only marginally varied with wind direction. On the other hand, the mean convective rain intensity of ~5 mm/h considerably varied across wind directions. The highest intensity was associated with SE circulations and the lowest with the NW circulations. Statistical data for each wind direction and rain type including standard deviation (SD) and standard error (SE) are summarized in Table 1.

Table 1. Summary of rain intensities (R) for wind directions in convective and stratiform rain.

Rain Type	Wind Direction	Duration (h)	Mean R (mm/h)	Median R (mm/h)	Standard Deviation (mm/h)	Standard Error (mm/h)
Convective	NE	82.5	4.51	3.65	4.90	0.070
	SE	50.7	6.23	4.83	6.41	0.116
	SW	645.6	5.11	3.72	6.09	0.031
	NW	538.1	4.33	3.36	4.89	0.027
	XX	269.6	5.80	4.54	5.79	0.046
Stratiform	NE	1191.9	0.79	0.50	0.79	0.003
	SE	486.2	0.80	0.46	0.92	0.005
	SW	5928.4	0.78	0.49	0.79	0.001
	NW	5740.0	0.83	0.54	0.80	0.001
	XX	3700.8	0.84	0.52	0.89	0.002

3.2. Rain Microstructure Variation With Rain Type and Wind Direction

Stratiform rain had smaller drops and lower drop concentrations compared to convective rain (Figure 5). The average D_0 for stratiform rain was 0.77 mm compared to 1.24 mm in convective rain. Normalized drop concentration N_w in stratiform rain was around 2.24×10^4 mm^{-1} m^{-3}, while convective rain had an average of 1.4×10^4 mm^{-1} m^{-3}. The overall average of D_0 (0.81 mm) and N_w (2.17×10^4 mm^{-1} m^{-3}) were closer to the values of stratiform rain since most rain intervals were of the stratiform type. The clusters in the values of D_0 that appear in Figure 5 emerge from the combined

effect of the diameter range bins of the disdrometer measurements, and the logarithmic scales on the horizontal axis.
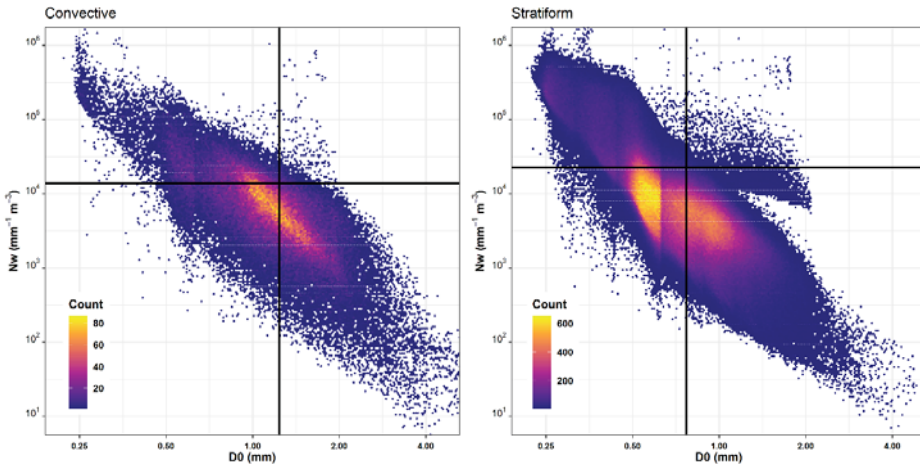
Figure 5. Scatter plot of D_0 and (N_W) for stratiform and convective rain. The vertical and horizontal lines represent the mean values of D_0 and N_W.

The distributions of D_0 and N_W values within each wind direction and rain type are illustrated in Figure 6. Similarly, the mean values of D_0 and N_W for different ranges of rain intensity within each wind direction and rain type are provided in Figure 7.

For stratiform rain, westerly circulations had larger drops and lower drop concentrations compared to easterly circulations. Especially SW had the largest mean D_0 and the lowest N_W. Easterly circulations were clearly characterized by the smallest drops and the greatest N_W. The same pattern was present even when inspecting different classes of rain intensity within stratiform rain (Figure 7). With higher rain intensity, D_0 increased too while N_W decreased.

For convective rain, only few differences in the previously described patterns were obvious especially when examining the rain microstructure for different ranges of rain intensities. With the exception of SE which had a limited number of convective intervals compared to the remaining wind directions, the median diameter D_0 was still the largest in SW and the smallest in NE, while N_W was the largest NE and the smallest in Sw. XX and NW had similar N_W values but NW exhibited larger drop sizes on average. The wind direction SE did not show any consistent pattern across rain intensity ranges.

When fitting a gamma function to the average rain drop size distribution within each wind direction in stratiform rain (Figure 8), easterly circulations had relatively lower concentrations of drops with a D_0 larger than 1 mm compared to westerly circulations. On the other hand, westerly circulations, especially SW, had the lowest concentration of drops with D_0 less than 1 mm. In convective rain, northerly circulations exhibited higher proportion of small drops ($D_0 < 1$ mm) and a smaller proportion of large drops compared to southerly circulations. Fitting gamma distribution to rain microstructure was also performed event by event. An example of the fitting for individual events is presented in Figure A1, and the density plots of the gamma distribution parameters are provided in Figure A2.
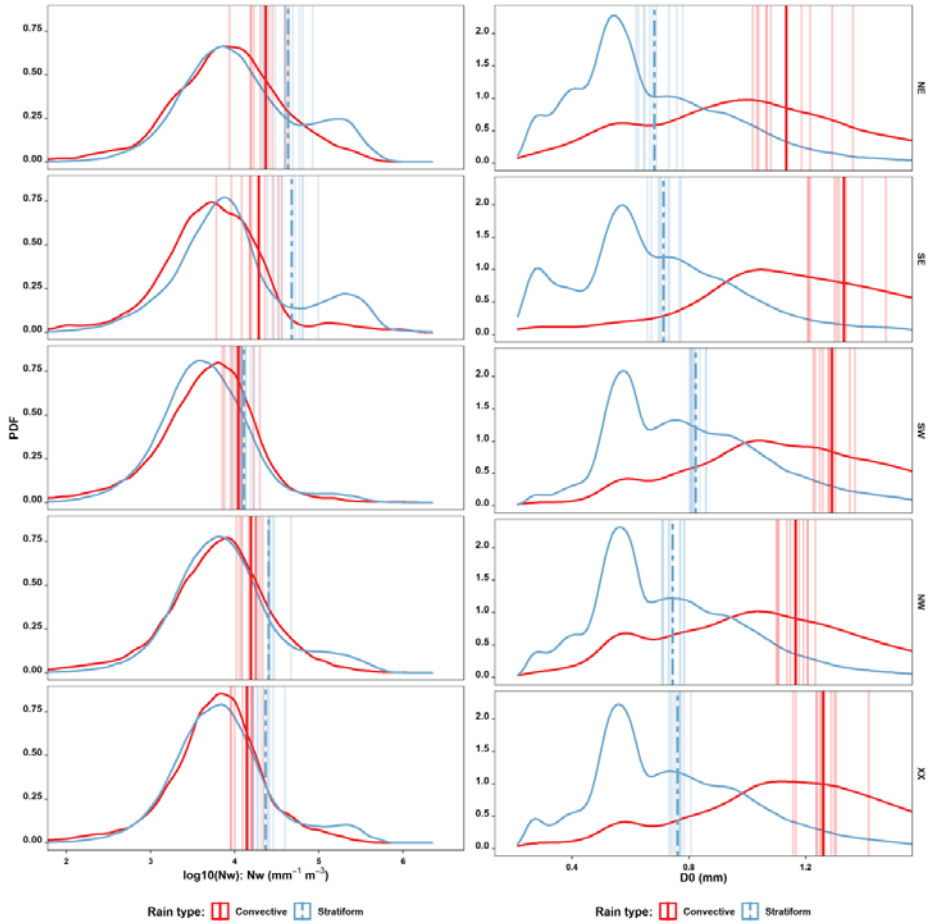
Figure 6. Probability density plots of log10(N_W) and D_0 for each rain type and wind direction. Vertical thick lines show parameters averaged over all locations for convective (solid red line) and stratiform (dashed thick blue) rain and light lines represent averages for individual locations.
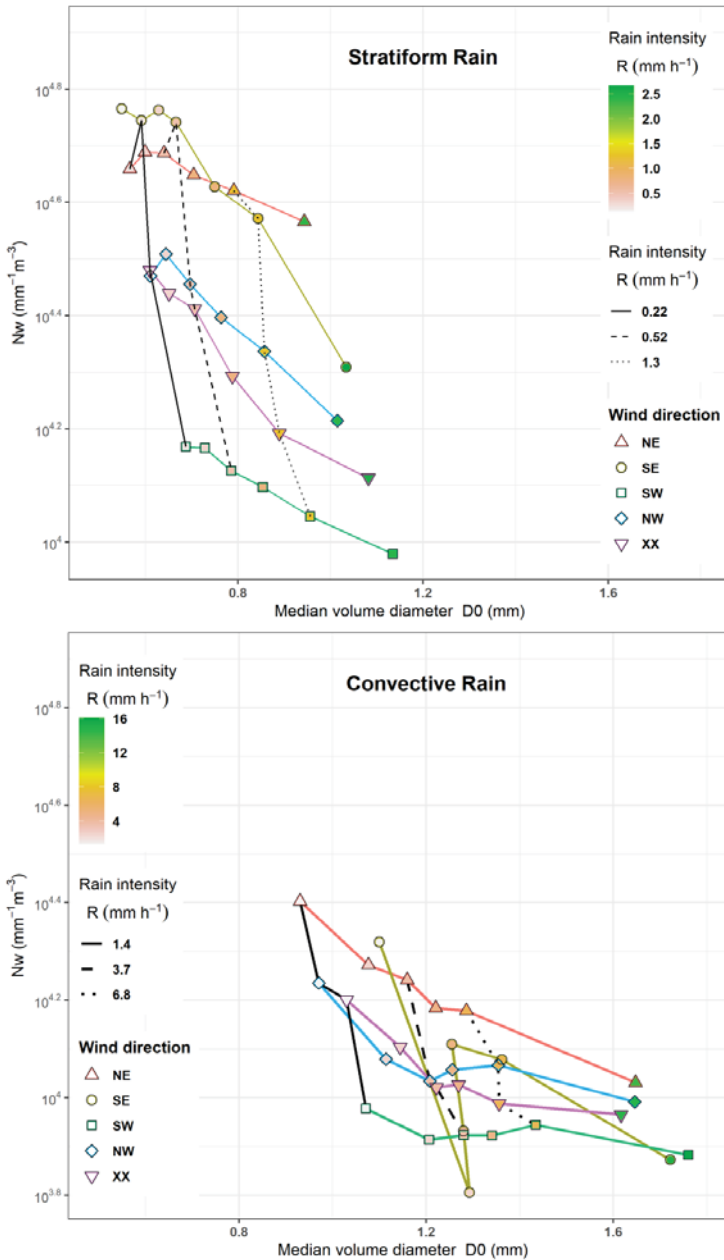
Figure 7. Rain microstructure for different rain intensities in stratiform and convective rain. Symbols on each colored line represent summary statistics for a wind direction. Each symbol represents the average median drop size D_0 and the normalized drop concentration for a rain intensity range. The intervals were chosen to represent six equal sample sizes and were colored by mean rain intensity. Selected symbols that correspond to equal rain intensity were connected with differently dashed black lines for comparison.
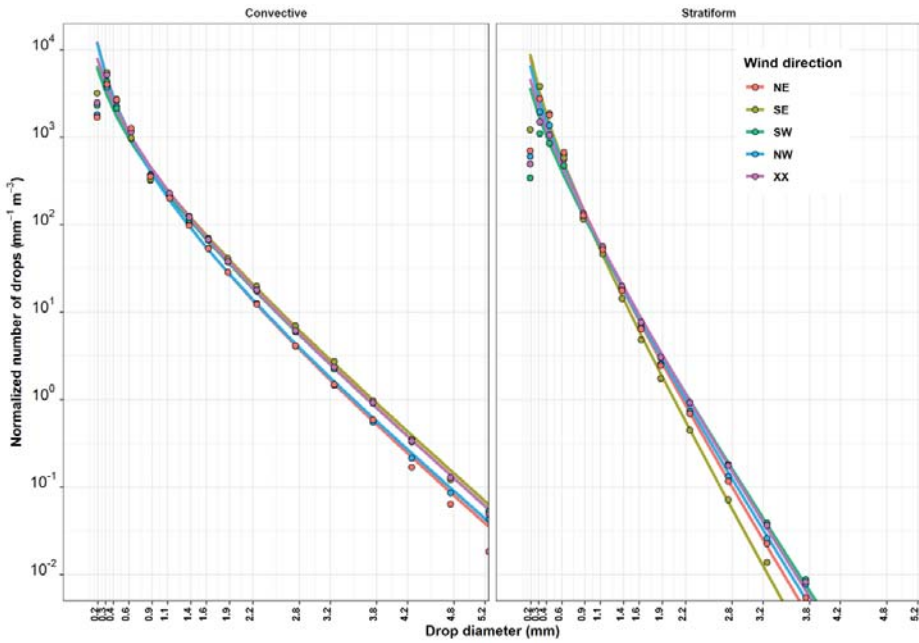
Figure 8. Raindrop concentration per millimeter diameter and cubic meter for each wind direction in stratiform rain and convective rain. Points represent the one-minute average concentrations for each diameter range colored by wind direction. Colored lines represent the corresponding gamma distribution fits.

3.3. Z–R Parameter Variation With Location, Rain Type and Wind Direction

To investigate the influence of rain microstructure variability per wind direction on the rain intensity retrieval equation Z–R, the values of A and b were obtained for 2449 events (see Section 2.4). A density plot of the R and dBZ values for all the 9914 h included in these events is provided in Figure A3. An example of the Z–R equation fitting for one event using two methods is provided in Figure A4.

The average value of the prefactor A was clearly larger in convective rain (309) than in stratiform rain (239), while the exponent b value was similar for both rain types (1.53). The values of A and b were averaged for each location (black points in Figure 9; Figure 10), for each wind direction (colored points in Figure 9; Figure 10), and for each combination of location and wind direction (colored stars in Figure 9; Figure 10) in order to demonstrate the variability of A and b with these factors.

In stratiform rain, the range of both mean A and b for each of the ten locations (the grey area in Figure 9) is comparable to the range of the average values for the wind directions (the red rectangle in Figure 9). However, A and b value are smaller in eastern circulation (NE, SE) compared to remaining general wind directions, and they are outside of the range associated with the spatial variability.

In convective rain, no clear pattern was detected for the average values of A and b associated with the five wind directions. The range of A and b values for the different locations is much larger than the range associated with the five general wind directions, indicating a larger spatial variability compared to the variability associated with general wind direction.

When averaging the values of A and b for each combination of location and wind direction, a greater range is observed. In the case of stratiform rain, the pattern of these values is comparable to the one observed for the five general wind directions; SW circulations have larger A values, easterly circulations have smaller A values, while XX and NW circulations fall closely in between.

The range of A and b values for each combination of the location and wind direction is larger in the case of convective rain. However, the small number of convective events needs to be considered in this case (see Table A1).
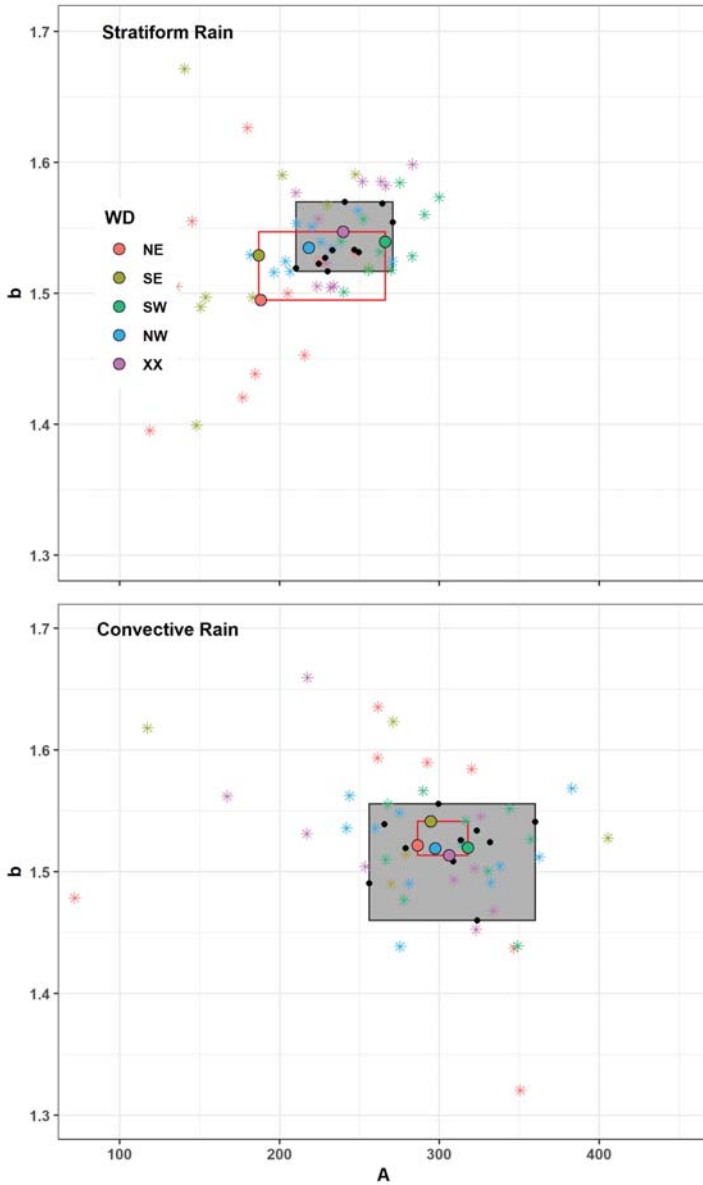
Figure 9. The parameters A and b of the radar rain intensity retrieval equation ($Z = AR^b$) in both rain types using the first method of fitting (Equations (13)–(19)). A and b values are averaged by location (black dots), wind direction (colored circles), and the combination of both (colored stars). The grey area represents the range of A and b for the ten locations. The red rectangle represents the range of A and b for the five general wind directions.
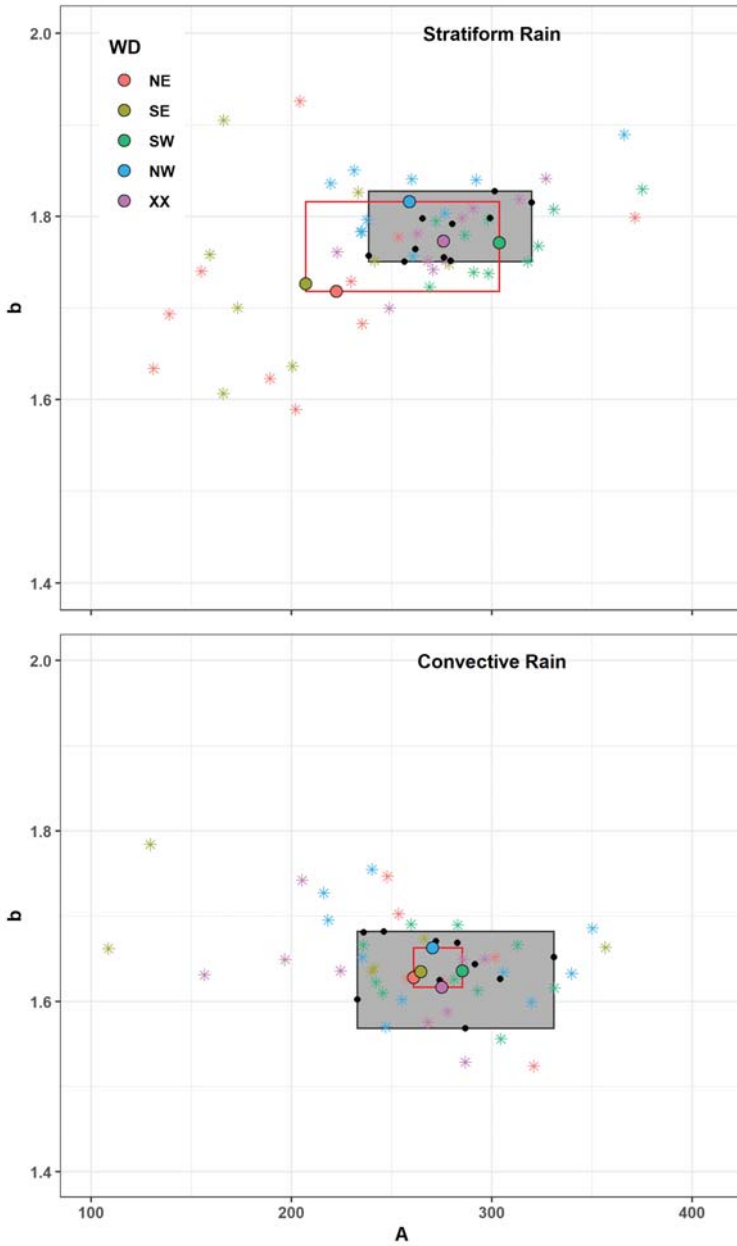
Figure 10. The parameters of the radar rain intensity retrieval equation ($Z = AR^b$) in both rain types using the second method of fitting (Equations (20)–(25)). A and b values are averaged by location (black dots), wind direction (colored circles), and the combination of both (colored stars). The grey area represents the range of A and b for the ten locations. The red rectangle represents the range of A and b for the five general wind directions.

4. Discussion

Our data indicate high frequency and high contribution of westerly and especially SW circulations to the rainy days over Bavaria, Germany. Easterly circulations have the least frequency and especially SE has the lowest share of rainy days. This is in agreement with the frequency of wind directions and proportions of rainy days of long-term studies for Germany for the period between 1995 and 2017 [28]. The high frequency and high contribution of westerly and southwesterly circulations to the number of rainy days is expected for this region since the main moisture flux is westerly [76].

Convection is responsible for 40% of rain amount in this region despite occupying only 10% of rain duration. Similar contributions of convective rain were reported for the Czech Republic [77] and in Switzerland [17]. Convective rain has typically higher rain rates and a distinct microstructure compared to stratiform rain. It is therefore essential to separate convective and stratiform rain prior to addressing rain microstructure, especially considering the variation in convective rain proportion with wind directions [17]. Southerly circulations generally have a higher proportion of convective rain compared to northerly circulations. A possible explanation is the strengthening and inhibition of convection and radiative cooling under different wind directions, which in turn has a major influence on the precipitation diurnal cycle over Germany [49]. Southerly circulations carry along warm air masses which intensify convection in the afternoon and inhibit radiative cooling in the early morning. Northerly circulations, in contrast, transport cold air masses, and therefore suppress convection and intensify radiative cooling.

Westerly circulations need special attention when addressing rain and microstructure, especially with the reported high contribution to rain duration and rain amount, and the expected increase in their frequency over Europe [78,79]. Westerly circulations are associated with larger rain drops than easterly circulations in stratiform rain, while easterly circulations have higher number of drops. This pattern is consistent for both stratiform and convective rain and across the ranges of rain intensity, except for SE circulations in convective rain, which was not well represented by data, accounting only for 0.6% of convective rain amount observed in this study.

Rain microstructure dependence on synoptic weather patterns has previously been reported for other locations in Europe. Northerly circulations in Leon, Spain, were associated with smaller drop sizes, while westerly and southerly circulations had larger rain drops [15]. This pattern was explained by the location of Leon to the south of the Cantabrian Mountains. Northerly circulation air masses precipitate prior to reaching Leon, leaving less humidity, lower rain intensities and smaller drops. Westerly and southerly circulations carry along higher humidity, leading to higher rain intensities and larger drops. For the Cévennes-Vivarais region in France, easterly circulations were associated with lower number of rain drops and larger drop size while most of the westerly circulations had the opposite traits [16]. The associations of rain microstructure with large-scale weather patterns observed in this and other studies are therefore not generally consistent, but region-specific. Different regions have different associated general air-mass characteristics, for example influenced by proximity to the sea or the presence of mountain massifs nearby. The origin of the air masses whether continental or maritime influences the rain microstructure and eventually influences the estimation of precipitation by radars [80,81]. Each class of wind direction used here has a mixture of both maritime and continental origins. It is however assumed that westerly circulations have a larger proportion of air masses with a maritime origin compared to easterly circulations.

The rain microstructure patterns in Bavaria have more in common with the patterns reported for Lausanne, Switzerland. Despite using different disdrometer types, schemes for rain type classification, and weather type classifications, and their different geographical locations in the Alps, easterly circulations were associated with higher number of drops per interval and smaller drop size compared to westerly circulations at both sites [17]. A plausible explanation for this is the variation of humidity and aerosol content in air masses between these wind direction clusters. Aerosols are particularly abundant in air masses which pass over Russia and Eastern Europe, especially over heavy industrialized areas [82,83]. These aerosols act as cloud condensation nuclei [84]. High availability of

cloud condensation nuclei increases the number of rain drops in the case of stratiform rain, increases the size of drops in local convection, but has no significant influence on rain microstructure in organized convection [85].

Differences in the load of cloud condensation nuclei under different circulations seem to be a plausible explanation for the rain microstructure differences observed in this study, especially in stratiform rain. The abundance of cloud condensation nuclei in easterly circulations in comparison with westerly circulations leads to higher number of rain drops. This in combination with the high (low) available humidity in westerly (easterly) circulations results in a larger (smaller) size of rain drops, respectively. For convective rain, easterly circulations comprise two wind directions, NE which has the smallest mean D_0, and SE which has the largest mean D_0. The larger size of raindrops in southerly circulations indicates the intensification of convection when the warm air masses are transported from the south, whereas northerly circulations bring colder airmasses. The rain type classification method used in this study does not differentiate local and organized convection, which makes it impossible to thoroughly compare with the findings of Cecchini et al. [85].

Our results may be useful for radar-based quantitative precipitation estimates (QPE), since Jaffrain et al. [75] demonstrated that the variation of A and b values in the Z–R retrieval equation is an important factor which should be accounted for. In their case study of Lausanne, Switzerland, spatial subgrid variability of rain microstructure was observed, which considerably influenced the quality of the estimation of rain rate. Using the same dataset, Ghada et al. [17] showed that the variability of A and b was larger than the subgrid spatial variability (in an area less than 1 km^2) when weather types are considered. In our study, variation of rain microstructure parameters with wind directions in Bavaria led to significant variation in the values of Z–R parameters. The variations in the prefactor A and the exponent b by wind direction are of a similar magnitude as their spatial variations in the case of stratiform rain, but smaller than the spatial variations in the case of convective rain. The same patterns were obtained for the conventional and the alternative methods of Z–R parameters retrieval despite the absolute differences in the values of A and b. These small differences occur because the conventional method is more sensitive to the large values of Z while the alternative method is more sensitive to the density of scatter points where R is below 2.5 mm/h [75]. This difference needs to be addressed in future studies to quantify the exact influence on the estimation of rain intensity by actual radar measurements. Alternatively, the least-rectangles linear regression could be applied as a middle-ground solution.

Assessing potential benefits of considering the variations in Z–R parameters, Jaffrain and Berne [75] concluded that the subgrid spatial variability in rain microstructure may account for errors in rain estimates between −2% and +15%. Variability due to large-scale weather patterns in Z–R parameters is likely to exceed their subgrid spatial variability [17], and based on our study, is comparable with the spatial variability of Z–R parameters in stratiform rain on a regional scale. Consequently, the potential for a significant improvement in rain estimation when accounting for rain microstructure variability by wind direction is expected to be high for radar quantitative precipitation estimates based only on radar reflectivity Z.

However, using only disdrometer data for this purpose would be insufficient because disdrometers provide a direct measurement of rain microstructure, from which R and Z are calculated. These values are accurate local measurements if we assume an accurate measurement of rain microstructure. The next logical research step would be a proper assessment of the improvement potential. This should include the integration of empirical data of radar-based rain intensity estimates validated by ground observations within the different rain types, locations, and large-scale wind directions, as well as a thorough rain type classification based on available instruments, especially considering the available network of dual polarization Doppler radars across Germany. Even for precipitation estimates based on a rain-gauge adjusted system as currently operated by DWD [86], improving the Z–R relation would likely have a positive impact in the final quality of the product.

5. Conclusions

This research demonstrated that rain microstructure varies significantly between weather types in both stratiform and convective rain. Easterly circulations had the highest drop concentration and the smallest drop size while westerly circulations were associated with large drops and low drop concentration. A plausible explanation for these differences is the high humidity content in westerly circulations and abundant cloud condensation nuclei concentration in easterly circulation. These finding offer potential new applications for radar-based quantitative precipitation estimates. Z–R parameters vary substantially with synoptic weather patterns effectively summarized by regional wind direction classes. This variation in Z–R parameters with wind direction approximates their station-to-station spatial variability for stratiform, but not for convective rain. We therefore conclude that building separate Z–R retrieval equation for regional wind direction classes should improve radar-based QPE, especially for stratiform rain events. This approach should be feasible for operational level forecasts, especially since daily large-scale weather types can be predicted with high accuracy several days in advance.

Appendix A

Table A1. Summary of events selected for the fitting of Gamma distribution and the two methods of R–Z parameters extraction (see Section 2.4).

Wind Direction	Rain Type	Duration (h)	# Events	Mean R (mm h^{-1})	Mean dBZ
SW	Convective	144.8	236	6.1	30.7
NW	Convective	85.7	131	5.0	29.3
XX	Convective	33.7	43	6.6	30.7
SE	Convective	11.1	10	6.5	31.6
NE	Convective	8.6	11	5.4	29.8
SW	Stratiform	3553.1	828	0.9	20.7
NW	Stratiform	3063.6	618	0.9	19.6
XX	Stratiform	2056.0	373	0.9	20.2
NE	Stratiform	598.8	140	0.9	18.5
SE	Stratiform	358.8	59	0.9	18.5
Total		9914.1	2449		
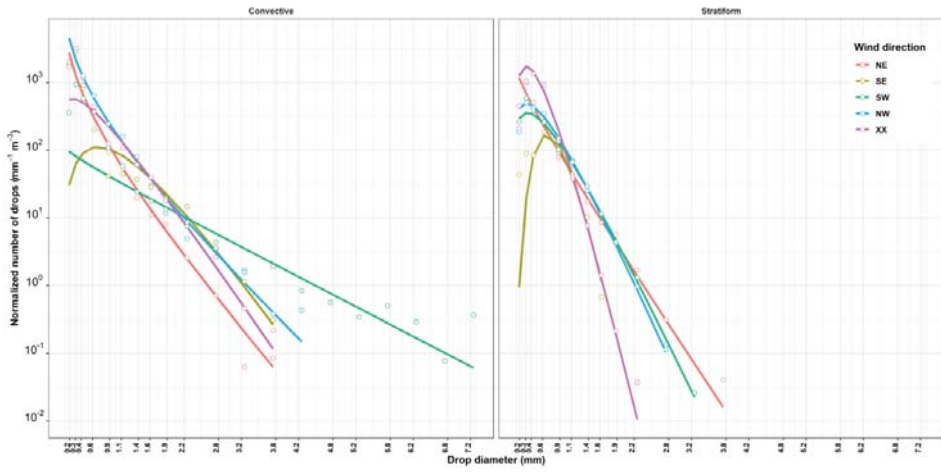
Figure A1. Raindrop concentration per millimeter and cubic meter for a selection of ten events; one event for each combination of wind direction and rain type. The points represent the event average one-minute concentrations for each diameter range colored by the relevant wind directions. The colored lines represent the fitted gamma distribution for these points.
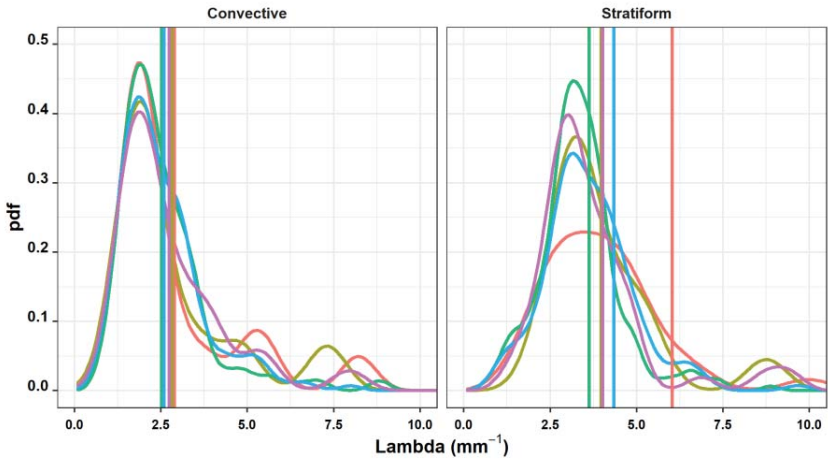
Figure A2. *Cont.*
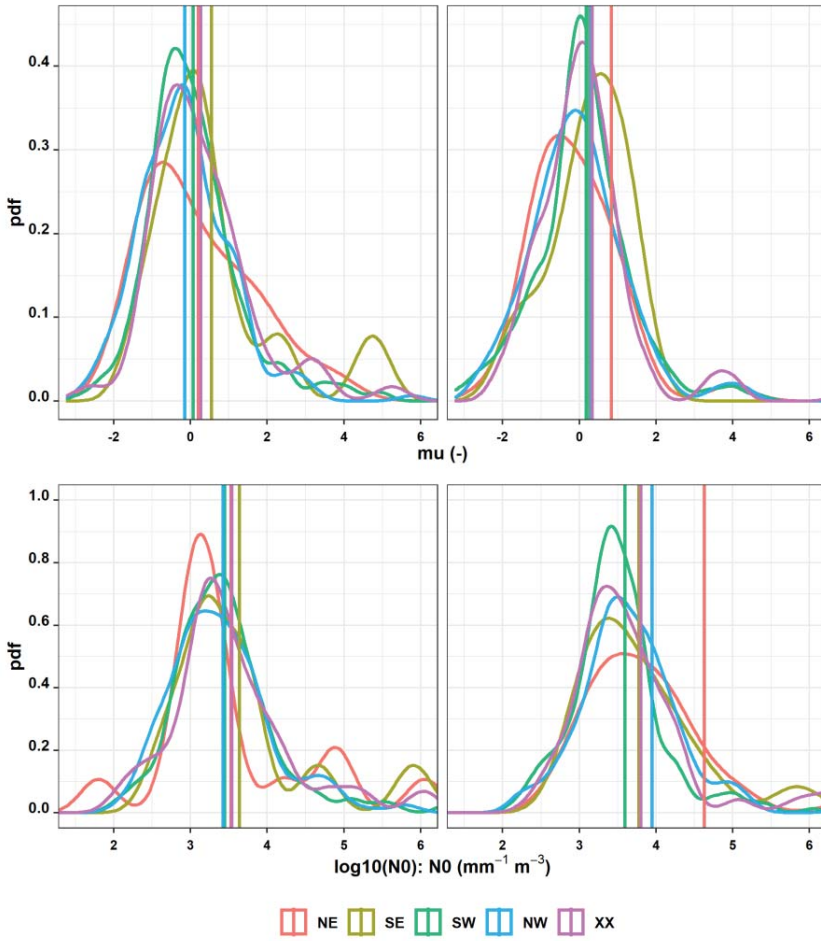
Figure A2. The probability density of the fitted gamma parameters in stratiform and convective rain colored by wind directions. The vertical lines represent the mean values of the three parameters based on a selection of 2449 events (see Section 2.4).
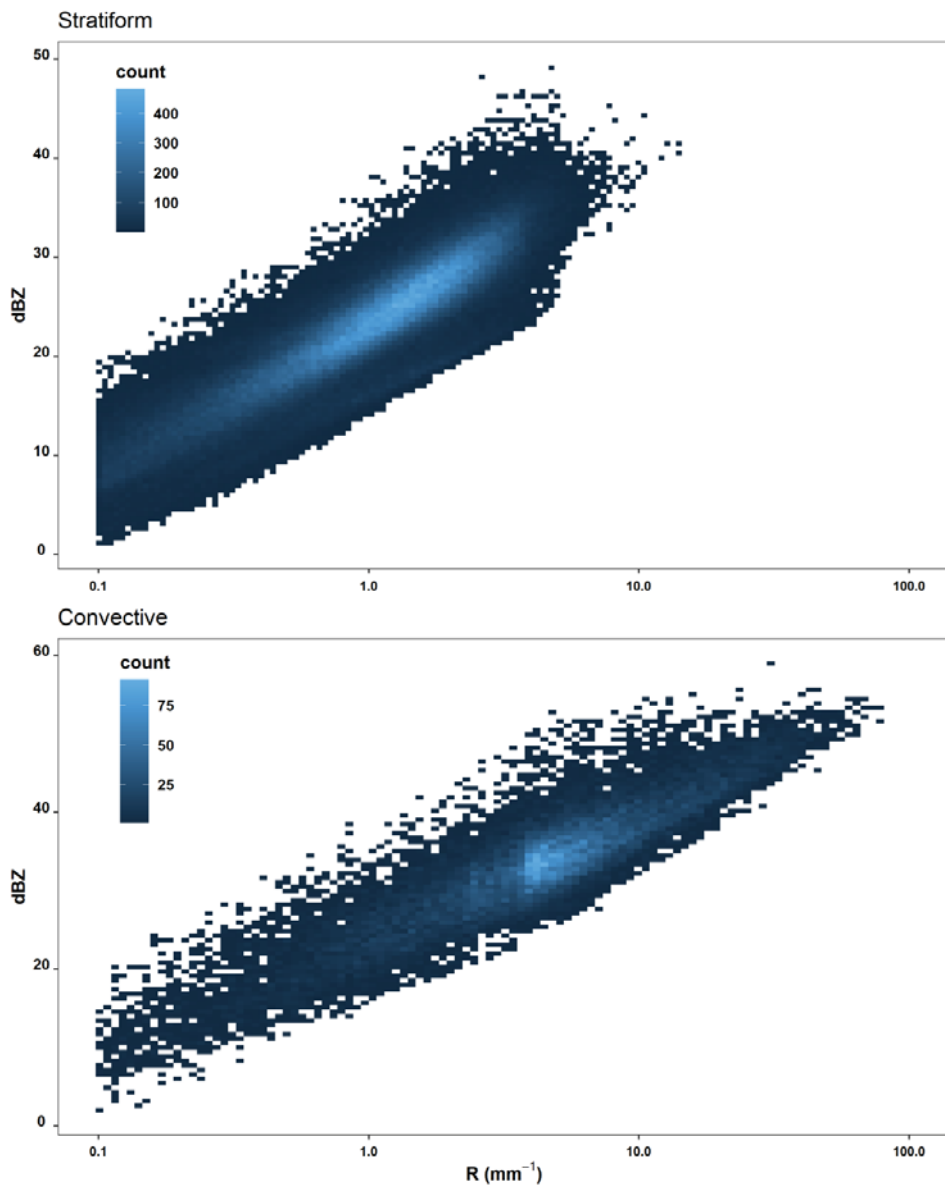
Figure A3. Density plot of reflectivity (dBZ) and rain intensity (R) for convective and stratiform rain. This plot includes the 9914 h of rain within the selected 2449 events (see Section 2.4).
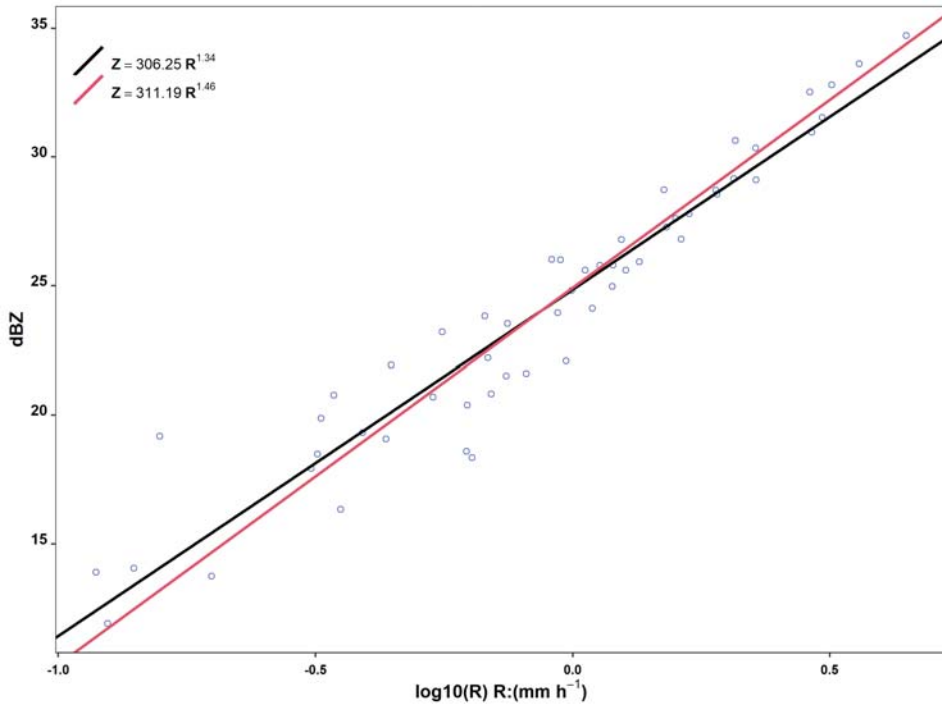
Figure A4. The fitted Z–R lines for one event (start: 2014-09-20 08:48:30, duration: 54 min, rain type: stratiform, wind direction: SW) using both the conventional method (black line) and the alternative method (red line).

References

1. Arulraj, M.; Barros, A.P. Improving quantitative precipitation estimates in mountainous regions by modelling low-level seeder-feeder interactions constrained by Global Precipitation Measurement Dual-frequency Precipitation Radar measurements. *Remote Sens. Environ.* **2019**, *231*, 111213. [CrossRef]
2. Steiner, M.; Smith, J.A.; Uijlenhoet, R. A Microphysical Interpretation of Radar Reflectivity–Rain Rate Relationships. *J. Atmos. Sci.* **2004**, *61*, 1114–1131. [CrossRef]
3. Thompson, E.J.; Rutledge, S.A.; Dolan, B.; Thurai, M. Drop Size Distributions and Radar Observations of Convective and Stratiform Rain over the Equatorial Indian and West Pacific Oceans. *J. Atmos. Sci.* **2015**, *72*, 4091–4125. [CrossRef]
4. Ryzhkov, A.V.; Zrnić, D.S. *Radar Polarimetry for Weather Observations*; Springer: Cham, Switzerland, 2019; ISBN 978-3-030-05092-4.
5. Steiner, M.; Smith, J.A. Convective versus stratiform rainfall: An ice-microphysical and kinematic conceptual model. *Atmos. Res.* **1998**, *47–48*, 317–326. [CrossRef]
6. Iacobellis, S.F.; McFarquhar, G.M.; Mitchell, D.L.; Somerville, R.C.J. The Sensitivity of Radiative Fluxes to Parameterized Cloud Microphysics. *J. Clim.* **2003**, *16*, 2979–2996. [CrossRef]
7. Jameson, A.R.; Larsen, M.L.; Kostinski, A.B. Disdrometer Network Observations of Finescale Spatial–Temporal Clustering in Rain. *J. Atmos. Sci.* **2015**, *72*, 1648–1666. [CrossRef]
8. Jaffrain, J.; Studzinski, A.; Berne, A. A network of disdrometers to quantify the small-scale variability of the raindrop size distribution. *Water Resour. Res.* **2011**, *47*, 2673. [CrossRef]
9. Doelling, I.G.; Joss, J.; Riedl, J. Systematic variations of Z–R-relationships from drop size distributions measured in northern Germany during seven years. *Atmos. Res.* **1998**, *47–48*, 635–649. [CrossRef]

10. Das, S.; Maitra, A. Characterization of tropical precipitation using drop size distribution and rain rate-radar reflectivity relation. *Appl. Clim.* **2018**, *132*, 275–286. [CrossRef]
11. Gatlin, P.N.; Thurai, M.; Bringi, V.N.; Petersen, W.; Wolff, D.; Tokay, A.; Carey, L.; Wingo, M. Searching for Large Raindrops: A Global Summary of Two-Dimensional Video Disdrometer Observations. *J. Appl. Meteor. Clim.* **2015**, *54*, 1069–1089. [CrossRef]
12. Dolan, B.; Fuchs, B.; Rutledge, S.A.; Barnes, E.A.; Thompson, E.J. Primary Modes of Global Drop Size Distributions. *J. Atmos. Sci.* **2018**, *75*, 1453–1476. [CrossRef]
13. Wen, L.; Zhao, K.; Wang, M.; Zhang, G. Seasonal Variations of Observed Raindrop Size Distribution in East China. *Adv. Atmos. Sci.* **2019**, *36*, 346–362. [CrossRef]
14. Niu, S.; Jia, X.; Sang, J.; Liu, X.; Lu, C.; Liu, Y. Distributions of Raindrop Sizes and Fall Velocities in a Semiarid Plateau Climate: Convective versus Stratiform Rains. *J. Appl. Meteor. Clim.* **2010**, *49*, 632–645. [CrossRef]
15. Fernandez-Raga, M.; Castro, A.; Marcos, E.; Palencia, C.; Fraile, R. Weather types and rainfall microstructure in Leon, Spain. *Int. J. Clim.* **2017**, *37*, 1834–1842. [CrossRef]
16. Hachani, S.; Boudevillain, B.; Delrieu, G.; Bargaoui, Z. Drop Size Distribution Climatology in Cévennes-Vivarais Region, France. *Atmosphere* **2017**, *8*, 233. [CrossRef]
17. Ghada, W.; Buras, A.; Lüpke, M.; Schunk, C.; Menzel, A. Rain Microstructure Parameters Vary with Large-Scale Weather Conditions in Lausanne, Switzerland. *Remote Sens.* **2018**, *10*, 811. [CrossRef]
18. Cerro, C.; Codina, B.; Bech, J.; Lorente, J. Modeling Raindrop Size Distribution and Z (R) Relations in the Western Mediterranean Area. *J. Appl. Meteor.* **1997**, *36*, 1470–1479. [CrossRef]
19. Thurai, M.; Gatlin, P.N.; Bringi, V.N. Separating stratiform and convective rain types based on the drop size distribution characteristics using 2D video disdrometer data. *Atmos. Res.* **2016**, *169*, 416–423. [CrossRef]
20. Munchak, S.J.; Kummerow, C.D.; Elsaesser, G. Relationships between the Raindrop Size Distribution and Properties of the Environment and Clouds Inferred from TRMM. *J. Clim.* **2012**, *25*, 2963–2978. [CrossRef]
21. Rosenfeld, D.; Ulbrich, C.W. Cloud Microphysical Properties, Processes, and Rainfall Estimation Opportunities. *Meteorol. Monogr.* **2003**, *30*, 237. [CrossRef]
22. Villermaux, E.; Bossa, B. Single-drop fragmentation determines size distribution of raindrops. *Nat. Phys.* **2009**, *5*, 697–702. [CrossRef]
23. Tokay, A.; Short, D.A. Evidence from Tropical Raindrop Spectra of the Origin of Rain from Stratiform versus Convective Clouds. *J. Appl. Meteor.* **1996**, *35*, 355–371. [CrossRef]
24. Caracciolo, C.; Prodi, F.; Battaglia, A.; Porcu', F. Analysis of the moments and parameters of a gamma DSD to infer precipitation properties: A convective stratiform discrimination algorithm. *Atmos. Res.* **2006**, *80*, 165–186. [CrossRef]
25. Caracciolo, C.; Porcù, F.; Prodi, F. Precipitation classification at mid-latitudes in terms of drop size distribution parameters. *Adv. Geosci.* **2008**, *16*, 11–17. [CrossRef]
26. Bringi, V.N.; Williams, C.R.; Thurai, M.; May, P.T. Using Dual-Polarized Radar and Dual-Frequency Profiler for DSD Characterization: A Case Study from Darwin, Australia. *J. Atmos. Ocean. Technol.* **2009**, *26*, 2107–2122. [CrossRef]
27. Bukovčić, P.; Zrnić, D.; Zhang, G. Convective–stratiform separation using video disdrometer observations in central Oklahoma—The Bayesian approach. *Atmos. Res.* **2015**, *155*, 176–191. [CrossRef]
28. Ghada, W.; Estrella, N.; Menzel, A. Machine Learning Approach to Classify Rain Type Based on Thies Disdrometers and Cloud Observations. *Atmosphere* **2019**, *10*, 251. [CrossRef]
29. Baur, F.; Hess, P.; Nagel, H. Kalender der grosswetterlagen Europas 1881–1939. *Bad Hombg.* **1944**, 35.
30. Steirou, E.; Gerlitz, L.; Apel, H.; Merz, B. Links between large-scale circulation patterns and streamflow in Central Europe: A review. *J. Hydrol.* **2017**, *549*, 484–500. [CrossRef]
31. Petrow, T.; Merz, B.; Lindenschmidt, K.-E.; Thieken, A.H. Aspects of seasonality and flood generating circulation patterns in a mountainous catchment in south-eastern Germany. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1455–1468. [CrossRef]
32. Nied, M.; Pardowitz, T.; Nissen, K.; Ulbrich, U.; Hundecha, Y.; Merz, B. On the relationship between hydro-meteorological patterns and flood types. *J. Hydrol.* **2014**, *519*, 3249–3262. [CrossRef]
33. Hofstätter, M.; Lexer, A.; Homann, M.; Blöschl, G. Large-scale heavy precipitation over central Europe and the role of atmospheric cyclone track types. *Int. J. Clim.* **2018**, *38*, e497–e517. [CrossRef] [PubMed]
34. Nikolopoulos, E.I.; Borga, M.; Marra, F.; Crema, S.; Marchi, L. Debris flows in the eastern Italian Alps: Seasonality and atmospheric circulation patterns. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 647–656. [CrossRef]

35. Kassomenos, P. Synoptic circulation control on wild fire occurrence. *Phys. Chem. Earth Parts A B C* **2010**, *35*, 544–552. [CrossRef]

36. Wastl, C.; Schunk, C.; Lüpke, M.; Cocca, G.; Conedera, M.; Valese, E.; Menzel, A. Large-scale weather types, forest fire danger, and wildfire occurrence in the Alps. *Agric. For. Meteorol.* **2013**, *168*, 15–25. [CrossRef]

37. Russo, A.; Trigo, R.M.; Martins, H.; Mendes, M.T. NO2, PM10 and O3 urban concentrations and its association with circulation weather types in Portugal. *Atmos. Environ.* **2014**, *89*, 768–785. [CrossRef]

38. Nidzgorska-Lencewicz, J.; Czarnecka, M. Winter weather conditions vs. air quality in Tricity, Poland. *Appl. Clim.* **2015**, *119*, 611–627. [CrossRef]

39. Grundström, M.; Dahl, Å.; Ou, T.; Chen, D.; Pleijel, H. The relationship between birch pollen, air pollution and weather types and their effect on antihistamine purchase in two Swedish cities. *Aerobiologia* **2017**, *33*, 457–471. [CrossRef]

40. Ramos, A.M.; Barriopedro, D.; Dutra, E. Circulation weather types as a tool in atmospheric, climate, and environmental research. *Front. Environ. Sci.* **2015**, *3*, 44. [CrossRef]

41. Huth, R.; Beck, C.; Philipp, A.; Demuzere, M.; Ustrnul, Z.; Cahynová, M.; Kyselý, J.; Tveito, O.E. Classifications of atmospheric circulation patterns: Recent advances and applications. *Ann. N. Y. Acad. Sci.* **2008**, *1146*, 105–152. [CrossRef]

42. Buishand, T.A.; Brandsma, T. Comparison of circulation classification schemes for predicting temperature and precipitation in The Netherlands. *Int. J. Clim.* **1997**, *17*, 875–889. [CrossRef]

43. Cortesi, N.; Trigo, R.M.; Gonzalez-Hidalgo, J.C.; Ramos, A.M. Modelling monthly precipitation with circulation weather types for a dense network of stations over Iberia. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 665–678. [CrossRef]

44. Broderick, C.; Fealy, R. An analysis of the synoptic and climatological applicability of circulation type classifications for Ireland. *Int. J. Clim.* **2015**, *35*, 481–505. [CrossRef]

45. Huth, R.; Beck, C.; Kučerová, M. Synoptic-climatological evaluation of the classifications of atmospheric circulation patterns over Europe. *Int. J. Clim.* **2016**, *36*, 2710–2726. [CrossRef]

46. Vallorani, R.; Bartolini, G.; Betti, G.; Crisci, A.; Gozzini, B.; Grifoni, D.; Iannuccilli, M.; Messeri, A.; Messeri, G.; Morabito, M.; et al. Circulation type classifications for temperature and precipitation stratification in Italy. *Int. J. Clim.* **2018**, *38*, 915–931. [CrossRef]

47. Twardosz, R. A synoptic analysis of the diurnal cycle of thunderstorm precipitation in Kraków (Southern Poland). *Int. J. Clim.* **2009**, *26*. [CrossRef]

48. Mandapaka, P.V.; Germann, U.; Panziera, L. Diurnal cycle of precipitation over complex Alpine orography: Inferences from high-resolution radar observations. *Q. J. R. Meteorol. Soc.* **2013**, *139*, 1025–1046. [CrossRef]

49. Ghada, W.; Yuan, Y.; Wastl, C.; Estrella, N.; Menzel, A. Precipitation Diurnal Cycle in Germany Linked to Large-Scale Weather Circulations. *Atmosphere* **2019**, *10*, 545. [CrossRef]

50. Cassano, E.N.; Lynch, A.H.; Cassano, J.J.; Koslow, M.R. Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA. *Clim. Res.* **2006**, *30*, 83–97. [CrossRef]

51. Planchon, O.; Quénol, H.; Dupont, N.; Corgne, S. Application of the Hess-Brezowsky classification to the identification of weather patterns causing heavy winter rainfall in Brittany (France). *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 1161–1173. [CrossRef]

52. Cony, M.; Martin, L. Synoptic patterns that contribute to extremely hot days in Europe. Available online: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-62362010000400001 (accessed on 30 October 2020).

53. Nowosad, J.; Stach, A. Relation between extensive extreme precipitation in Poland and atmospheric circulation. *Quaest. Geogr.* **2014**, *33*, 115–129. [CrossRef]

54. Maheras, P.; Tolika, K.; Anagnostopoulou, C.; Makra, L.; Szpirosz, K.; Károssy, C. Relationship between mean and extreme precipitation and circulation types over Hungary. *Int. J. Clim.* **2018**, *38*, 4518–4532. [CrossRef]

55. Kirsch, B.; Clemens, M.; Ament, F. Stratiform and convective radar reflectivity–rain rate relationships and their potential to improve radar rainfall estimates. *J. Appl. Meteor. Clim.* **2019**, *58*, 2259–2271. [CrossRef]

56. Friedrich, K.; Kalina, E.A.; Masters, F.J.; Lopez, C.R. Drop-Size Distributions in Thunderstorms Measured by Optical Disdrometers during VORTEX2. *Mon. Wea. Rev.* **2013**, *141*, 1182–1203. [CrossRef]

57. Atlas, D.; Srivastava, R.C.; Sekhon, R.S. Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.* **1973**, *11*, 1. [CrossRef]

58. Gunn, R.; Kinzer, G.D. The terminal velocity of fall for water droplets in stagnant air. *J. Meteorol.* **1949**, *6*, 243–248. [CrossRef]

59. Chen, B.; Wang, J.; Gong, D. Raindrop Size Distribution in a Midlatitude Continental Squall Line Measured by Thies Optical Disdrometers over East China. *J. Appl. Meteor. Clim.* **2016**, *55*, 621–634. [CrossRef]

60. Angulo-Martínez, M.; Beguería, S.; Latorre, B.; Fernández-Raga, M. Comparison of precipitation measurements by OTT Parsivel 2 and Thies LPM optical disdrometers. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2811–2837. [CrossRef]

61. Deutscher Wetterdienst. Description of the method and the weather types. Available online: https://www.dwd.de/EN/ourservices/wetterlagenklassifikation/beschreibung.html;jsessionid= 5ED0BB104CB32D74A71ADC0C6DE03871.live21064?nn=495490&lsbId=520444 (accessed on 27 September 2019).

62. Deutscher Wetterdienst. Weather Type Classification Data. Available online: https://www.dwd.de/EN/ourservices/wetterlagenklassifikation/online_wlkdaten.txt?view=nasPublication&nn=495490 (accessed on 15 October 2019).

63. R Core Team. *R: A Language and Environment for Statistical Computing*; EEA: Vienna, Austria, 2019.

64. RStudio Team. *RStudio: Integrated Development Environment for R.*; PBC: Boston, MA, USA, 2018.

65. Kuhn, M. Caret: Classification and Regression Training. 2018. Available online: https://CRAN.R-project.org/package=caret (accessed on 30 October 2020).

66. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. *E1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071)*; TU: Wien, Austria, 2018; Available online: https://CRAN.R-project.org/package=e1071 (accessed on 30 October 2020).

67. Wickham, H. Reshaping Data with the reshape Package. *J. Stat. Softw.* **2007**, *21*, 1–20. [CrossRef]

68. Robert, J.H. Raster: Geographic Data Analysis and Modeling. 2017. Available online: https://CRAN.R-project.org/package=raster (accessed on 30 October 2020).

69. Ryan, M.H. Rmisc: Rmisc: Ryan Miscellaneous. 2013. Available online: https://CRAN.R-project.org/package=Rmisc (accessed on 30 October 2020).

70. Hadley, W. *Ggplot2. Elegrant Graphics for Data Analysis*, 2nd ed.; Springer: Cham, Switzerland, 2016; ISBN 978-3-319-24277-4.

71. South, A. Rnaturalearth: World Map Data from Natural Earth. 2017. Available online: https://CRAN.R-project.org/package=rnaturalearth (accessed on 30 October 2020).

72. Ulbrich, C.W. Natural Variations in the Analytical Form of the Raindrop Size Distribution. *J. Clim. Appl. Meteor.* **1983**, *22*, 1764–1775. [CrossRef]

73. Ulbrich, C.W.; Atlas, D. Rainfall Microphysics and Radar Properties: Analysis Methods for Drop Size Spectra. *J. Appl. Meteor.* **1998**, *37*, 912–923. [CrossRef]

74. Wetterdienst, D. Historical records of hourly cloud type in Germany. Available online: ftp://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/cloud_type/historical/ (accessed on 16 October 2019).

75. Jaffrain, J.; Berne, A. Influence of the Subgrid Variability of the Raindrop Size Distribution on Radar Rainfall Estimators. *J. Appl. Meteor. Clim.* **2012**, *51*, 780–785. [CrossRef]

76. van der Ent, R.J.; Savenije, H.H.G.; Schaefli, B.; Steele-Dunne, S.C. Origin and fate of atmospheric moisture over continents. *Water Resour. Res.* **2010**, *46*, 61. [CrossRef]

77. Rulfová, Z.; Kyselý, J. Trends of Convective and Stratiform Precipitation in the Czech Republic, 1982–2010. *Adv. Meteorol.* **2014**, *2014*, 1–11. [CrossRef]

78. Plavcová, E.; Kyselý, J. Projected evolution of circulation types and their temperatures over Central Europe in climate models. *Appl. Clim.* **2013**, *114*, 625–634. [CrossRef]

79. Stryhal, J.; Huth, R. Trends in winter circulation over the British Isles and central Europe in twenty-first century projections by 25 CMIP5 GCMs. *Clim. Dyn.* **2019**, *52*, 1063–1075. [CrossRef]

80. Das, S.; Chatterjee, C. Rain characterization based on maritime and continental origin at a tropical location. *J. Atmos. Sol. Terr. Phys.* **2018**, *173*, 109–118. [CrossRef]

81. Tenório, R.S.; Cristina da Silva Moraes, M.; Sauvageot, H. Raindrop Size Distribution and Radar Parameters in Coastal Tropical Rain Systems of Northeastern Brazil. *J. Appl. Meteor. Clim.* **2012**, *51*, 1960–1970. [CrossRef]

82. Birmili, W.; Wiedensohler, A.; Heintzenberg, J.; Lehmann, K. Atmospheric particle number size distribution in central Europe: Statistical relations to air masses and meteorology. *J. Geophys. Res.* **2001**, *106*, 32005–32018. [CrossRef]

83. Byčenkienė, S.; Plauškaitė, K.; Dudoitis, V.; Ulevicius, V. Urban background levels of particle number concentration and sources in Vilnius, Lithuania. *Atmos. Res.* **2014**, *143*, 279–292. [CrossRef]

84. Lohmann, U.; Feichter, J. Global indirect aerosol effects: A review. *Atmos. Chem. Phys.* **2005**, *5*, 715–737. [CrossRef]

85. Cecchini, M.A.; Machado, L.A.T.; Artaxo, P. Droplet Size Distributions as a function of rainy system type and Cloud Condensation Nuclei concentrations. *Atmos. Res.* **2014**, *143*, 301–312. [CrossRef]

86. Kreklow, J.; Tetzlaff, B.; Burkhard, B.; Kuhnt, G. Radar-Based Precipitation Climatology in Germany—Developments, Uncertainties and Potentials. *Atmosphere* **2020**, *11*, 217. [CrossRef]

Article

ST-CORAbico: A Spatiotemporal Object-Based Bias Correction Method for Storm Prediction Detected by Satellite

Miguel Laverde-Barajas [1,2,3,4,*,†], Gerald A. Corzo [1], Ate Poortinga [3,5], Farrukh Chishtie [3,5], Chinaporn Meechaiya [3,4], Susantha Jayasinghe [3,4], Peeranan Towashiraporn [3,4], Amanda Markert [6,7], David Saah [3,4,8], Lam Hung Son [9], Sothea Khem [9], Surajate Boonya-Aroonnet [10], Winai Chaowiwat [10], Remko Uijlenhoet [11,‡] and Dimitri P. Solomatine [1,2]

1   IHE-Delft Institute for Water Education, Westvest 7, 2611 AX Delft, The Netherlands; g.corzo@un-ihe.org (G.A.C.); d.solomatine@un-ihe.org (D.P.S.)
2   Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2628 CN Delft, The Netherlands
3   SERVIR-Mekong, SM Tower, 24th Floor, 979/69 Paholyothin Road, Samsen Nai Phayathai, Bangkok 10400, Thailand; apoortinga@sig-gis.com (A.P.); fchishtie@sig-gis.com (F.C.); chinaporn.m@adpc.net (C.M.); susantha@adpc.net (S.J.); peeranan@adpc.net (P.T.); dsaah@sig-gis.com (D.S.)
4   Asian Disaster Preparedness Center, SM Tower, 24th Floor, 979/69 Paholyothin Road, Samsen Nai Phayathai, Bangkok 10400, Thailand
5   Spatial Informatics Group, LLC, 2529 Yolanda Ct., Pleasanton, CA 94566, USA
6   Earth System Science Center, The University of Alabama in Huntsville, 320 Sparkman Drive, Huntsville, AL 35805, USA; amanda.m.weigel@nasa.gov
7   SERVIR Science Coordination Office, NASA Marshall Space Flight Center, 320 Sparkman Drive, Huntsville, AL 35805, USA
8   Geospatial Analysis Lab, University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117, USA
9   Mekong River Commission, Flood Management and Mitigation Programme, 576 National Road No. 2, Sangkat Chak Angre Krom, Khan Menachey, Phnom Penh 12353, Cambodia; son@mrcmekong.org (L.H.S.); khem@mrcmekong.org (K.S.)
10 Hydroinformatic Institute, 901 Ngam Wong Wan Road, Lat Yao, Chatuchak, Bangkok 10900, Thailand; surajate@hii.or.th (S.B.-A.); winai@hii.or.th (W.C.)
11 Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, Wageningen University, 6700 AA Wageningen, The Netherlands; remko.uijlenhoet@wur.nl
*   Correspondence: m.laverde@un-ihe.org
†   SM Tower, 24th Floor, 979/69 Paholyothin Road, Samsen Nai Phayathai, Bangkok 10400, Thailand.
‡   Current address: Delft University of Technology, 2628 CN Delft, The Netherlands.

Abstract: Advances in near real-time rainstorm prediction using remote sensing have offered important opportunities for effective disaster management. However, this information is subject to several sources of systematic errors that need to be corrected. Temporal and spatial characteristics of both satellite and in-situ data can be combined to enhance the quality of storm estimates. In this study, we present a spatiotemporal object-based method to bias correct two sources of systematic error in satellites: displacement and volume. The method, Spatiotemporal Contiguous Object-based Rainfall Analysis for Bias Correction (ST-CORAbico), uses the spatiotemporal rainfall analysis ST-CORA incorporated with a multivariate kernel density storm segmentation for describing the main storm event characteristics (duration, spatial extension, volume, maximum intensity, centroid). Displacement and volume are corrected by adjusting the spatiotemporal structure and the intensity distribution, respectively. ST-CORAbico was applied to correct the early version of the Integrated Multi-satellite Retrievals for the Global Precipitation Mission (GPM-IMERG) over the Lower Mekong basin in Thailand during the monsoon season from 2014 to 2017. The performance of ST-CORAbico

is compared against the Distribution Transformation (DT) and Gamma Quantile Mapping (GQM) probabilistic methods. A total of 120 storm events identified over the study area were classified into short and long-lived storms by using a k-means cluster analysis method. Examples for both storm event types describe the error reduction due to location and magnitude by ST-CORAbico. The results showed that the displacement and magnitude correction made by ST-CORAbico considerably reduced RMSE and bias of GPM-IMERG. In both storm event types, this method showed a lower impact on the spatial correlation of the storm event. In comparison with DT and GQM, ST-CORAbico showed a superior performance, outperforming both approaches. This spatiotemporal bias correction method offers a new approach to enhance the accuracy of satellite-derived information for near real-time estimation of storm events.

1. Introduction

Rainfall is a key component in the hydrological cycle and the primary source of freshwater in many regions. However, climatic extremes, such as floods and landslides that are caused by heavy rainfall events, pose a great threat to communities causing loss of life and damage to properties [1,2]. Flood monitoring and water management applications require a high accuracy representation of rainfall in extreme conditions [3]. This is of particular importance in tropical monsoonal climates, such as the Mekong basin region, where convective storm events are localised in space and time.

Traditionally, rain gauges and ground-based weather radar networks have provided the most reliable precipitation data at the catchment scale [4]. However, in many areas around the world, these data are either scarce or not available. An irregular distribution of these ground-based observation stations makes it difficult to discern spatiotemporal features of convective storms and their associated rainfall fields. Therefore, satellite-based remote sensing measurements have become an important source of rainfall data e.g., [5–7]. These satellite-based measurements enable monitoring of storm events at a quasi-global scale in near real-time. In comparison to other remote systems, such as ground-based weather radar, satellite sensor observations provide a wider coverage, being able to acquire data on storm systems in regions where topographic variations limit or obstruct weather radars [8].

Satellite-based precipitation estimates are derived by combining visible to long-wave infrared (VIS/IR) sensors from the Geosynchronous Earth Orbit (GEO) satellite with Passive Microwave (PMW) sensors from the Low Earth Orbit (LEO) satellites. VIS/IR sensors are relevant to measure albedo and cloud top temperature with a high temporal and spatial resolution [9,10]. On the other hand, PMW sensors can penetrate clouds for measuring thermal emissions, which are attenuated by raindrops with a 3-h interval [11,12]. Currently available and commonly used Satellite-based Precipitation Products (SPP) include the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [13], the NOAA Climate Prediction Center MORPHing technique (CMORPH [14]), the Multi-satellite Precipitation Analysis from the Tropical Rainfall Measurement Mission (TMPA) [15,16], and the Integrated Multi-satellitE Retrievals from the Global Precipitation Measurement (GPM-IMERG) [17]. Sun et al. [18] details a comprehensive review of the main global available satellite-based precipitation datasets.

Despite advances in the field of remote sensing, SPP information is subject to several systematic and random errors that require correction e.g., [19–21]. A wide range of bias correction methodologies has been developed to improve the performance of SPP, leveraging ground-based observations. Several examples include linear scaling, local intensity scaling, the power and distribution transformation methods, and Gamma Quantile mapping e.g., [8,22,23]. These methods all

adjust SPP as function of rainfall intensity values, ignoring important systematic errors, such as those that are caused by displacement and timing.

In the field of weather forecasting, displacement error in storm prediction has been taken using spatial verification methods into account [24–26]. These "nontraditional" methods do not rely on point-to-point matches between the observed and estimated fields for avoiding double penalties (e.g., rainfall estimated but not observed and vice versa) that are commonly found in traditional approaches. Methods can be broadly grouped into neighbour or fuzzy [27,28], scale separation [29–31], object-based e.g., [24–26], and field transformation [32,33]. The first two categories can be described as spatial filtering methods, in which the verification statistics are evaluated at coarser resolutions to provide information about the scale of the performance. Object-based and field transformation are considered as displacement verification methods when estimated rainfall fields, defined as an object, are spatially manipulated (displacement, rotation, scaling, etc.) to try to fit the observed value.

Several studies have used spatial verification methods to analyse and correct systematic error of SPP based on the characteristics of matched storm objects, such as location, rotation, intensity, and shape [34–36]. For instance, Demaria et al. [37] used the object-based method, Contiguous Rainfall Analysis (CRA, Ebert and McBride [24]), to correct the location error of CMORPH, PERSIANN, and the TMPA datasets over the Plata basin. Recently, Le Coz et al. [38] used the field transformation method, called Feature Calibration and Alignment technique (FCA), to correct the error due to location in the GPM-IMERG late version over Sub-Saharan Africa. These methods have been useful for correcting the displacement errors when the grid resolution is high and the storm event is small, while preserving the higher spatial variability of SPP storm. However, these methodologies are constrained by the two-dimensional analysis of the storm event.

The spatiotemporal analysis can provide a much deeper analysis on aspects of the entire life-cycle of the storm event, including time span, speed, evolution, among others. In the literature, error analysis using spatiotemporal approaches has been useful to evaluate the performance of several spatial rainfall products. For example, Ref. [39,40] used the Object-Based Diagnostic Evaluation time-domain (MODE-TD) that was proposed by Bullock [41] to evaluate the convection-allowing forecast from the Weather Forecast Model over the United States. Recently, Laverde-Barajas et al. [42] used the Spatiotemporal Contiguous Rainfall Analysis (ST-CORA; Laverde-Barajas et al. [43]) in the Southeast region of Brazil for analysing the error composition of the CMORPH SPP and evaluated the individual hydrological response of two systematic error sources: location and magnitude. This study demonstrated the importance of spatial and temporal storm characteristics to analyse the main systematic error sources in SPP.

Spatiotemporal storm analysis incorporated into bias correction methods is key to reduce several sources of systematic error in SPP. In this study, we present a spatiotemporal object-based bias correction method to reduce several systematic errors in storm events estimated by satellite. The method, called Spatiotemporal Contiguous Object-based Rainfall Analysis for Bias Correction (ST-CORAbico), uses the main storm characteristics of satellite and observed events detected by the ST-CORA method to remove errors due to displacement in space and time and volume. This method is evaluated over the lower Mekong Basin in Thailand to correct several storm event types in the Integrated Multi-satellitE Retrievals for GPM (GPM-IMERG) early version during the monsoon season from 2014 to 2017. The performance of ST-CORAbico is compared against two widely used probabilistic methods—Distribution Transformation and Gamma Quantile Mapping. This manuscript is organised, as follows: Section 2 describes the study and the rainfall data-sets; Section 3 details the methodology of ST-CORAbico; Sections 4 and 5 contain the results and associated discussion; and finally, Section 6 presents the conclusions and future work.

2. Study Area and Data

2.1. Study Area and Period

The study was conducted in Isan, in the northeastern region of Thailand (Figure 1). This region is part of the Mekong-river basin, with an average discharge of 475 km^3/year. The Mekong river basin covers a total drainage area of 795.000 km^2 and it is characteristically divided into the upper and lower basins. The Lower Mekong river in Thailand is located in the Lower Mekong basin, covering 23% of the total drainage area (184.000 km^2) and representing 36% of the country (Figure 1).
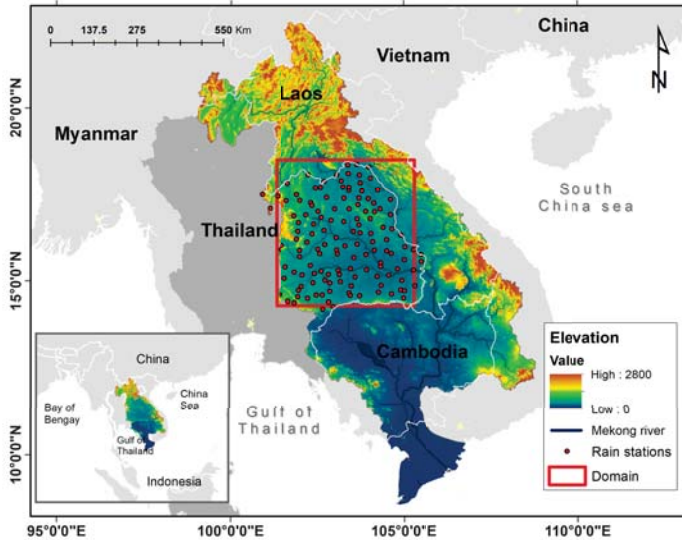
Figure 1. Digital elevation of the Lower Mekong Basin in Thailand. Red rectangle corresponds to the study area in the Isan region. Red dots represent the hourly rain gauge stations from www.thaiwater.net. Blue line represents the Lower Mekong river network.

Heavy rainfall often occurs in the Lower Mekong river basin during the Asian monsoon period between June to October. In this period, the covariability of the Indian summer monsoon and East Asian summer monsoon impacts the Lower Mekong basin, causing high convective storm events that are the main trigger for flash floods and landslides [44,45]. We focused on the wet months for the years 2014 to 2017, as these disasters are mostly occurring during the wet season.

2.2. Satellite-based Precipitation Data

ST-CORAbico was used to correct storm events detected by the Integrated Multi-satellitE Retrievals for GPM (GPM-IMERG). GPM-IMERG is a high-resolution global precipitation product produced by [17] from the NASA/JAXA Global Precipitation Measurement (GPM) mission. The product combines information from multiple infrared, passive-microwave, and satellite-radar sensors to provide rainfall estimations at 0.1-degree spatial resolution every half-hour. GPM-IMERG computes Early, Late, and Final runs. The first two runs are near real-time versions of IMERG and are available at six hours and 18 hours latency, respectively. In the Early version, rainfall estimations are propagated forward while the Late has both forward and backward propagation allowing the incorporation of climatological gauge data. The final version is obtained three months after the measurements. In this run GPM-IMERG ingests the monthly rainfall analysis from the Global Precipitation Climatology Centre (GPCC; Schneider et al. [46]). This version is used for scientific

purposes, as it is considered to be the most reliable version [17]. In this study, we evaluated the early run of GPM-IMERG at a half-hourly temporal resolution. This version was selected, given it is the lowest latency data product available, a crucial aspect for operational applications.

2.3. In-Situ Ground-Based Rainfall Observations

Hourly rainfall data during the monsoon season from 2014 to 2017 were obtained from a dense network of 138 rain gauge stations from the Thailand Integrated Water Resource Management System, operated by the Hydro-Informatics Institute (HII) in Thailand (Figure 1). Data were quality controlled by mirroring the density distribution functions of neighbouring stations to remove outlier values considered as noise. These observations were then further interpolated using the Ordinary Kriging interpolator from Golden Software [47] at 0.1 degrees for each hour to match the spatial and temporal resolution of the GPM-IMERG data. This method was selected due to the moderated topographic conditions of the study area and density of the rain gauge measurements. Raingauge data were interpolated using an exponential variogram with a sill of 1.14 $(\mathrm{mm}^2/\mathrm{h}^2)$, a range of 4.4 km and a nugget of 0.66 $(\mathrm{mm}^2/\mathrm{h}^2)$. It must be noted that interpolation methods for rainfall data are subject to uncertainty e.g., [48,49]; however, in our case, a dense and optimal rain gauge distribution can reduce the level of uncertainty from the interpolation method [50].

3. Methodology

ST-CORAbico was developed in order to analyse the spatiotemporal characteristics of storm events and bias correct the main sources of systematic error in satellites. Figure 2 shows the methodology of ST-CORAbico. In this section, we describe the elements for storm analysis and bias correction in ST-CORAbico.
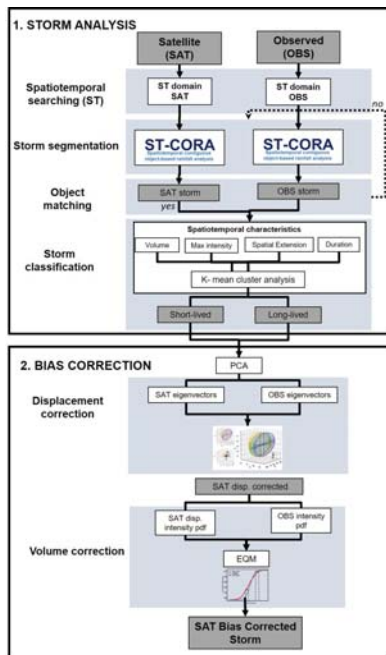
Figure 2. Diagram of the Spatiotemporal Contiguous Object-based Rainfall Analysis for bias correction (ST-CORAbico) method. Grey boxes represent the input and output products while white boxes describe the methodological process for storm analysis and bias correction components.

3.1. Storm Analysis

In the storm analysis, ST-CORAbico uses ST-CORA to analyse the spatiotemporal characteristics of the storm events observed and detected by satellites. This process requires the definition of the spatial and temporal domain in order to reduce the computational time of ST-CORA. We applied a spatiotemporal searching algorithm to predetermine the region of analysis in ST-CORA. This algorithm uses the spatial searching algorithm concept that was proposed by Guttman [51] to index areas with rainfall information in both datasets. The indexing is made in a two-dimensional space compressing the latitude and longitude dimensions using a maximum intensity value as a reference. Once the spatiotemporal domain is defined, we use ST-CORA in the observed and SPP dataset to identify storms in the rainfall data. In this study, ST-CORA incorporates a multivariate kernel density function for storm segmentation.

3.1.1. Storm Segmentation Using the Spatiotemporal Object-Based Rainfall Analysis with Multivariate Kernel Density Segmentation

ST-CORA was applied to analyse the spatiotemporal characteristics of storm events at the catchment scale (duration, spatial extent, magnitude, and centroid). This method enables the feature extraction of different storm event types, classified based on hydrometeorological criteria. ST-CORA uses a multidimensional connected labelling component algorithm to associate connected voxels in space and time (a volume generalisation of pixels) into a disjoint object labelled with a unique classifier. This operation is built upon binary information that was created by voxels, considered to be 'effective rainfall'. Effective rainfall voxels $S_{[x,y,t]}$ are defined according to rainfall voxels $R_{x,y,t}$ above the rainfall intensity threshold IT, as:

$$S_{[x,y,t]} := \begin{cases} 1, & \text{if } R_{x,y,t} \geq \text{IT.} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

where, IT is defined by the user and $S_{[x,y,t]}$ is defined in terms of 1 = "true" or 0 ="false". In this study, we used $IT = 1$mm/h to define effective rainfall [52]. Once binary voxels are created, the connected labelling component algorithm scans all voxels in a neighbour system (from top to bottom and left to right), assigning preliminary labels to $S_{[x,y,t]}$, as follows:

$$c(S_{[x,y,t]}) = \left\{ N_{[x,y,t]} \in \alpha s : S_{CR} = S_N \right\} \tag{2}$$

where, $c(S_{[x,y,t]})$ is a preliminary label, S_{CR}, S_N are properties of the voxel $S_{[x,y,t]}$ and its neighbours $N_{[x,y,t]}$, respectively, while αs is the neighbour system in space and time. The labelling process $c(S_{[x,y,t]})$ is repeated to resolve equivalence classes of the spatiotemporal object.

Bethel et al. [53] found that object segmentation, while using image thresholding, such as the connected component labelling method, has limitations for edge detection in data with unknown topology. In the original ST-CORA, a size-filtering algorithm and morphological closing method are incorporated in order to remove both small noisy objects and a false merging effect, respectively. However, this process is based on a binary object not taking into account the intensity value of voxels. To overcome this limitation, we have incorporated a Multivariate Kernel Density Estimation (KDE) approach to segment rainfall objects when considering their four dimensions. This method assumes a non-parametric probability density distribution technique for d-dimensional data. Notably, KDE has been widely used in many fields for image detection and object tracking, e.g., [54–58]. Multivariate kernel density is estimated at point x from a random sample $X_1, X_2, ...X_n$ from a density function, f,

$$\widehat{f_K}(x) = \frac{1}{n} \sum_{i=1}^{n} K_h (x - x_i) \tag{3}$$

where K corresponds to the kernel function and h is the bandwidth matrix. Choosing the bandwidth matrix can be restricted to a class of positive diagonal matrices [59]. In the literature, there are several

bandwidth selection methods for kernel density estimation [59,60]. For this approach, we use the normal reference rule-of-thumb proposed by Henderson and Parmeter [61]. This method estimates the bandwidth while assuming that the density distribution function follows a Gaussian distribution.

The process of edge detection using KDE is based on the Edge Detection by Density method that was developed by Pereira et al. [55]. This process evaluates the multivariate density distribution of the density of a four-dimensional (4D) rainfall object (Figure 3), and segments the object based on the density threshold, u. This threshold identifies the storm edges that are lower than a probability percentage. This parameter is calculated by analysing the relationship between threshold delineation and the connected intensity value. We found that the 25th distribution percentile for u threshold showed good results for storm segmentation over the Lower Mekong Basin, especially for intense storm events, which are characteristic of monsoon environments.

Rainfall Objects are considered to be storm based on the Critical Mass Threshold (CMT), which is defined as the minimum volume of rainfall (km^3) necessary to be considered as an extreme event [62]. The value of CMT is calculated locally based on the sensitivity between the spatial extent and the total object volume [37,63]. In this analysis, we also incorporated the sensitivity of CMT to the maximum intensity of the storm in order to evaluate the response of intense storm events in the study area. Based on the sensitivity analysis of those parameters, we selected a CMT of 0.01 km^3 for storm events with a maximum intensity greater than 10 mm/h. In the study area, these events correspond to rainfall objects bigger than 2000 km^2.
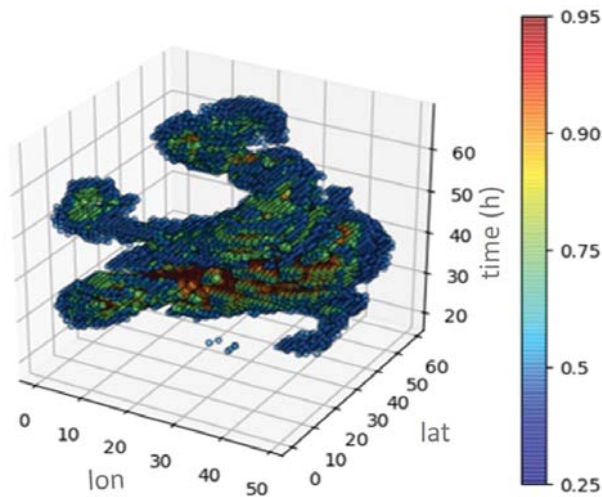
Figure 3. Multivariable kernel density of a storm object in space and time. Example for the storm event 2014-07.

3.1.2. Matching Process

As a result of ST-CORA, multiple storm events are identified in both observed and satellite data sets. However, it is necessary to determine the observed and estimated storm matches. We used the Intersection-over-Union measure (IoU) in order to evaluate the level of similarity between predicted and observed data. IoU is defined as the ratio between the size of the intersection and the union of both objects. This is represented by the following equation:

$$IoU = \frac{TP}{FP + TP + FN} \tag{4}$$

where, *TP* represents true positives and *FP* and *FN* are false positives and negatives, respectively. The selection of the intersection rate value determines the level of matching between objects. If the values are too low, multiple objects will be indexed with the same object. On the other hand, high values indicate that the object does not have any match. We found the *IoU* value to be 30 percent in the selected study area, which is a good balance for matching the observed and satellite storm events.

3.1.3. Storm Classification

Once all of the storm events are identified, ST-CORA classifies storm events into two types: small convective systems with a short duration (short-lived) and long duration systems extended over large areas (long-lived) [64,65]. We used an unsupervised K-means cluster analysis method to classify short- and long-lived storm events based on the four main storm characteristics (duration, spatial extent, maximum intensity, and total volume). This method divides *n* observations into *k* clusters in which each observation is a member of the cluster that minimises the objective function *J*, as follows:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \| X_{(x,y,z,c)i}^{j} - C_j \|^2 \tag{5}$$

where, *X* is the storm with dimensions *x*, *y*, *z*, *c* corresponding to the storm characteristics duration, spatial extent, maximum intensity and total volume, respectively. *C* is the centroid of the cluster *k* and the absolute number represents the minimum Euclidean distance to *C*.

3.2. Bias Correction

Bias correction is the second component of the ST-CORAbico method. This component is based on the systematic error source extraction for SPP that was proposed by Laverde-Barajas et al. [42]. Based on the error decomposition for storm estimation defined by Ebert and McBride [24], satellite error is composed of systematic and aleatory errors due to displacement, volume, and pattern, as:

$$E_{total} = E_{displacement} + E_{volume} + E_{pattern} \tag{6}$$

where displacement and volume represent the systematic errors and pattern is the aleatory error calculated as follows:

$$E_{displacement} = E_{total} - E_{shifted} \tag{7}$$

$$E_{volume} = E_{total} - E_{magnitude} \tag{8}$$

$$E_{pattern} = E_{shifted} - E_{volume} \tag{9}$$

In Equations (7) and (8), location is the main source of error due to displacement, while the magnitude is the corresponding source of error for volume. Using the error subtraction from Laverde-Barajas et al. [42], ST-CORAbico corrects displacement and volume error using the following process:

Displacement Correction

Displacement correction corresponds to removing the shifting effect of the estimated storm. In this step, the Principal Component Analysis method (PCA, Johnson and Hebert [66]) is used in order to obtain the weighted centroid and orthogonal variables (eigenvectors, eigenspace) of the SPP storm and reference data (Figure 4a). Once the geometric properties of the objects are obtained, the weighted centroid is matched and the object is rotated accordingly to fit the eigenvectors of the reference storm data.
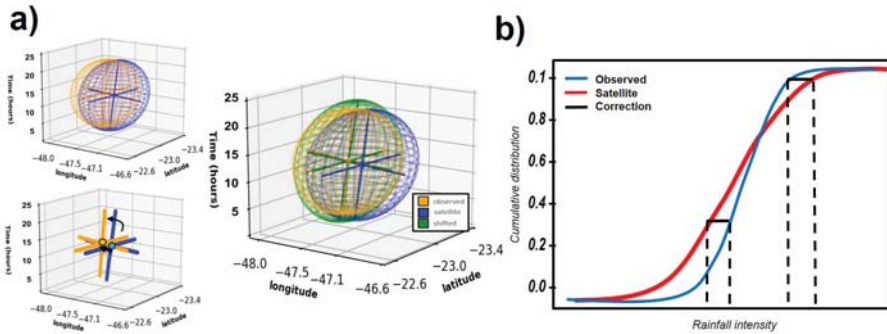
Figure 4. ST-CORAbico systematic error subtraction. (**a**) In error correction, storm centroid and eigenvectors derived from principal component of GPM-IMERG are fitted to the observed event; and, (**b**) magnitude subtraction using both satellite and observed empirical distribution functions, with respect to intensity.

Volume Correction

Volume correction corresponds to the subtraction of the magnitude source of error of the SPP. Using the statistical Empirical Quantile Method (EQM) [67], magnitude error is subtracted by adjusting all moments of the empirical cumulative distribution functions ($ecdfs$) of the SPP in terms of intensity, with respect to the reference data (Figure 4b). EQM builds the $ecdf$ for the observed ($ecdf_{obs}$) and the satellite ($ecdf_{sat}$) while using the intensity storm distribution I_s, as:

$$EQM = ecdf_{obs}^{-1}(ecdf_{sat}(I_s)) \qquad (10)$$

3.3. Evaluation of ST-CORAbico

The evaluation was done by comparing the bias-corrected results with two widely used probabilistic bias correction methods— the Distribution Transformation (DT) method and the Gamma Quantile Mapping (GQM). The DT method was originally developed for the statistical downscaling of climate model data [68]. The method corrects the mean and difference in variation of the SPP by matching the satellite and the observed distribution based on Equation (11):

$$DT = (SAT(t) - \mu_{sat})DT_\tau + \tau_{sat} \times DT_\mu \qquad (11)$$

where, μ and τ are the mean and standard deviation of the observed and satellite, respectively. DT_μ and DT_τ are the mean and standard deviation ratio between the observed and satellite data at time t.

The Gamma Quantile Mapping method uses the same methodology as the Empirical Quantile mapping method (10), based on the assumption that both observed *OBS* and satellite *SAT* intensity follows a gamma distribution [69]. DT and GQM are implemented for each time-step in order to correct the storm event. The bias correction performance is evaluated based on three widely used error metrics: the Root Mean Square Error (Equation (12)) for evaluating the magnitude error, the bias level (Equation (13)) to evaluate the systematic bias, and the correlation coefficient (Equation (14)) in order to analyse the linear correlation between the observed and the bias-corrected storm event.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(OBS_i - SAT_i)^2} \qquad (12)$$

$$Bias = \frac{\sum\limits_{i=1}^{N} (SAT_i - OBS_i)}{\sum\limits_{i=1}^{N} (OBS_i)} \tag{13}$$

$$r = \frac{\sum\limits_{i=1}^{N} (SAT_i - \overline{SAT})(OBS_i - \overline{OBS})}{\sqrt{\sum\limits_{i=1}^{N} (SAT_i - \overline{SAT})^2} \sqrt{\sum\limits_{i=1}^{N} (OBS_i - \overline{OBS})^2}} \tag{14}$$

where, *OBS* represents the rainfall values of the reference rain gauge data and *SAT* are the satellite and the bias-corrected storm obtained with each method.

4. Results

4.1. Storm Analysis

We identified 120 storm events observed and estimated by GPM-IMERG at an hourly scale for the 2014–2017 monsoon seasons. Figure 5 shows the scatter plot of the main storm characteristics (total volume, duration, spatial extent, and maximum intensity) and classification between short- and long-lived storm events using the k-means cluster analysis. For all events, 68 storms (56%) were classified as short-lived storms while 52 (44%) of storms were classified as long-lived events. Short-lived events had a duration that ranged between three and 17 h, with a maximum spatial extent of 42 thousand km^2. Long-lived events had a duration ranging between 18 and 31 h and covered between 54 and 110 thousand km^2. In terms of total volume and maximum intensity, short-lived events have a total volume of up to 0.15 km^3 with low and intense storms ranging from 3 to 82 mm/h. On the other hand, long-lived storms are comprised of medium and high-intensity events with a total volume ranging from 0.27 to 0.65 km^3. Table 1 describes the observed storm characteristics for short and long-lived event types.
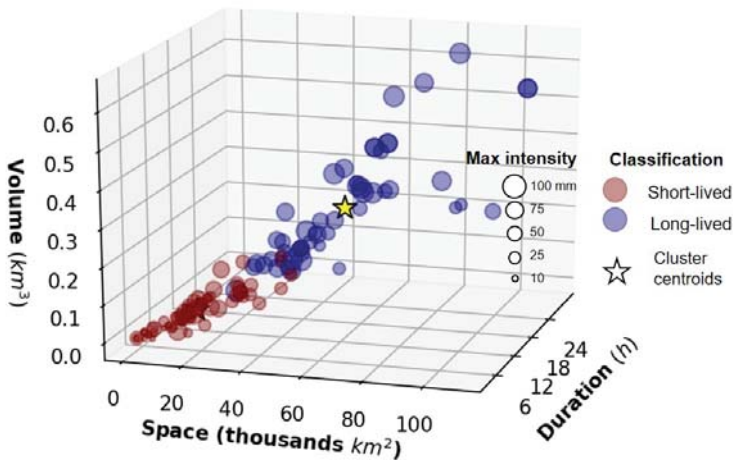
Figure 5. Short- and long-lived cluster analysis classification for observed events during monsoon season 2014–2017.

Table 1. Storm characteristics for short- and long-lived event types.

Storm Type	Statistics	Duration (h)	Spatial Extent (km^2)	Maximum Intensity (mm/h)	Total Volume (km^3)
Short-lived storm	mean	9	15,097	33.0	0.04
	min	3	1900	3.6	0.01
	max	17	42,300	82.0	0.15
long-lived storm	mean	18	54,400	71.4	0.27
	min	10	24,300	31.6	0.07
	max	31	110,600	100.0	0.64

4.2. Results for Bias Correction

We selected a short-lived and a long-lived storm in order to describe the workflow for displacement and volume correction made by ST-CORAbico. Figures 6 and 7 present the bias correction steps for each storm event type. Panel (a) shows the spatial distribution of the observed and satellite events as well as the bias-corrected satellite storm events that were obtained from the correction of location and magnitude errors in ST-CORAbico. Panel (b) describes the displacement and volume corrections. Panel (c) presents the four-dimensional (4D) spatiotemporal evolution of the observed and satellite as well as the bias-corrected storm (time in the z-axis). Panel (d) shows the bias and RMSE statistics as well the scatter and correlation between the observed storm and original and bias-corrected satellite events.

Both examples (Figures 6 and 7) show the importance of bias correction. In both the short and long-lived event scenarios, GPM-IMERG had a longer duration with a larger footprint. However, the long-lived event presented a better spatial agreement than the short-lived event. In terms of magnitude, GPM-IMERG considerably overestimated the total volume and rainfall intensity of the storm. Overall, the performance of GPM-IMERG shows a positive bias and high RMSE, mostly being caused by an excess of rainfall. The correlation coefficients for short- and long-lived event scenarios were 0.7 and 0.5, respectively.
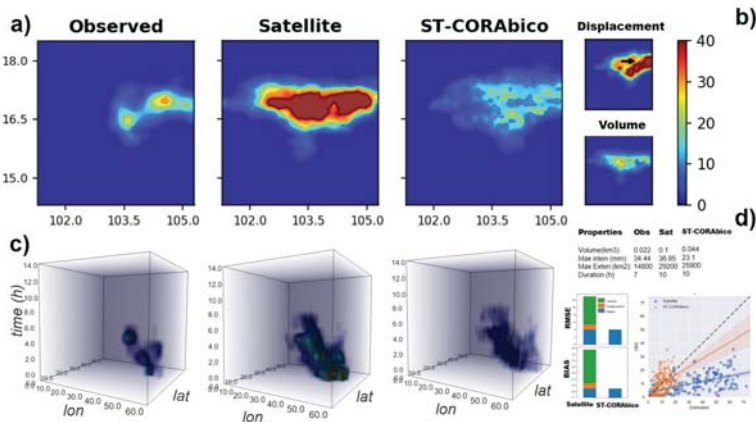
Figure 6. Performance of ST-CORAbico for a short-lived storm event (2014-08-27). (**a**) total events for observed, satellite and ST-CORAbico; (**b**) volume, displacement correction maps; (**c**) four-dimensional (4D) spatiotemporal evolution (lat, lon, time, intensity); and, (**d**) bias, RMSE statistics, and scatter and correlation between observed and estimated rainfall values.

The corrections in displacement and volume made by ST-CORAbico displayed notable changes in the satellite storm structure. In both scenarios, RMSE and bias were mostly reduced by correction due to volume, contributing 40 to 60% of the RMSE reduction and around 70% of the total bias reduction for both events. Displacement correction had an important impact on the reorientation of the satellite storm. The individual correction contributed to 5% of the RMSE correction and 10% reduction of the

total bias for the short-lived event. In the case of the long-lived scenario, displacement correction contributed almost 15% of the RMSE reduction and 20% of the total bias reduction. In terms of the correlation coefficient, displacement and volume correction that were made by ST-CORAbico showed a marginal impact on the spatial correlation for the short-lived events. For the long-lived scenario, this did not impact the spatial correlation.
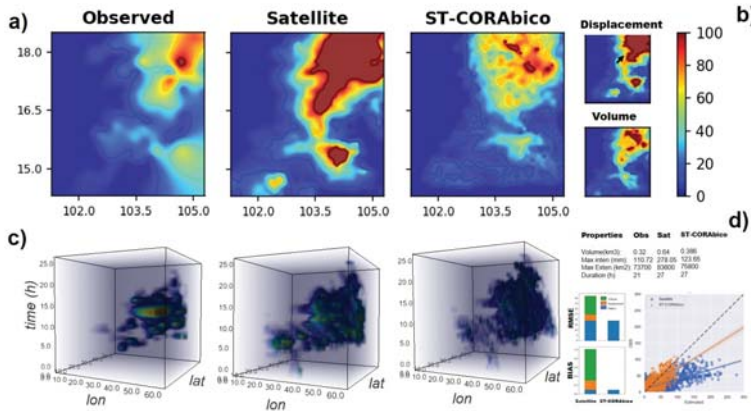
Figure 7. Performance of ST-CORAbico for a long-lived storm event (2014-07-21). (**a**) total events for observed, satellite and ST-CORAbico; (**b**) volume, displacement correction maps; (**c**) 4D spatiotemporal evolution (lat, long, time, intensity); and, (**d**) bias, RMSE statistics and scatter and correlation between observed and estimated rainfall values.

Figure 8 presents the performance of ST-CORAbico for short- and long-lived storm events. This figure describes the density distribution of RMSE (a–b), bias (c–d), and correlation coefficient (e–f) of the short- and long-lived storms estimated by GPM-IMERG, ST-CORAbico, and the individual corrections due to displacement and volume. It was found that ST-CORAbico has a smaller error distribution in RMSE and bias for short- and long-lived storm events when compared with the original GPM-IMERG. This error reduction is mostly caused by the correction due to volume. Displacement correction was an important factor in reducing the bias, especially for long-lived storm events. The results from the correlation coefficient showed that ST-CORAbico had a marginal effect on the spatial correlation of the storm event. Overall, it was found that ST-CORAbico considerably reduced the systematic error of GPM-IMERG.

4.3. Model Comparison

ST-CORAbico was compared with the Distribution Transformation method (DT) and the Gamma Quantile Mapping (GQM) method. Using the short- and long-lived storm scenarios that are presented above, Figure 9 presents the spatial differences and linear correlation between the total observed storms and the bias-corrected events obtained by ST-CORAbico, DT, and GQM. The results for both storm event scenarios showed that ST-CORAbico had the lowest spatial difference among the evaluated methods. For the short-lived storm scenario, ST-CORAbico displayed the highest correlation coefficient (r: 0.41) and the lowest RMSE and bias (RMSE: 4.05 mm; bias: 0.74) when compared with DT (r: 0.40; RMSE: 5.4 mm; bias: 1.17); and, GQM (r: 0.39 RMSE: 6.09mm and bias: 1.5). In the case of the long-lived storm, ST-CORAbico and DT showed a notable error reduction in contrast to the GQM method that showed the biggest differences. For this storm scenario, ST-CORAbico had the best performance (r: 0.71 RMSE: 18.02 mm; bias: 0.09), followed by DT (r: 0.68, RMSE: 23.0 mm; bias: 0.32), and finally GQM (r: 0.62, RMSE: 43.77 mm; bias: 0.97).
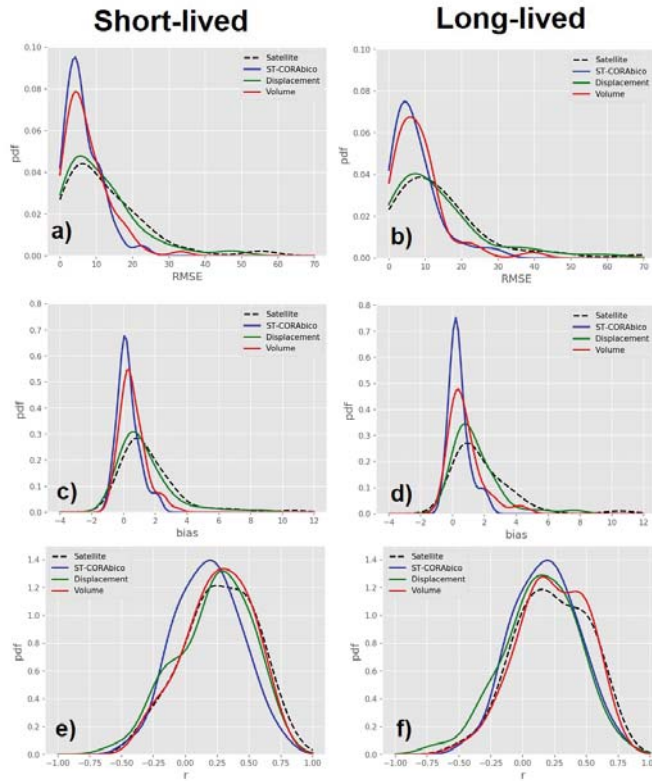
Figure 8. Satellite and bias corrected error distribution for short and long-lived events during monsoon seasons 2014–2017: (**a**,**b**) RMSE; (**c**,**d**) bias; and, (**e**,**f**) correlation coefficient.
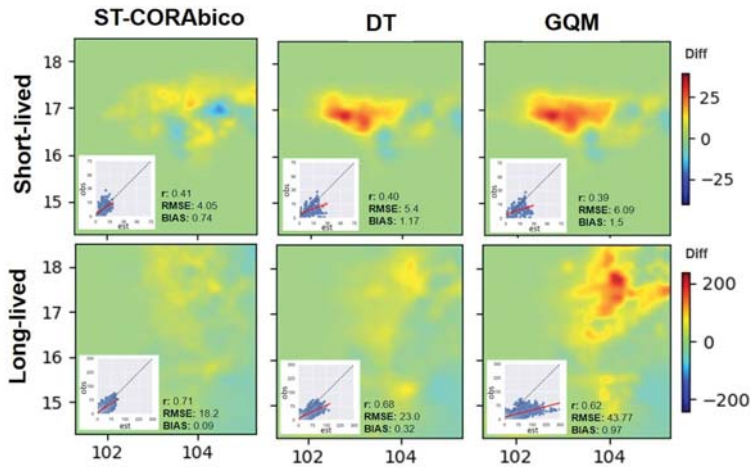
Figure 9. Comparison between ST-CORAbico vs Distribution Transformation (DT) and Gamma Quantile Mapping (GQM).

Figure 10 presents the comparison between ST-CORAbico, DT, and GQM for short- and long-lived storm events. The boxplots show the distributions of the RMSE (Figure 10a,b), the bias (Figure 10c,d), and the correlation coefficient (Figure 10d,f) between the 25% and 75% percentiles for the original GPM-IMERG and the different bias correction methods. The dots represent the individual error for each storm event. In comparison with the two probabilistic methods, we found that ST-CORAbico consistently had the lowest RMSE as well as the lowest bias for both short- and long-lived storm events. ST-CORAbico and DT had a lower impact on the correlation coefficient, especially for short-lived events.
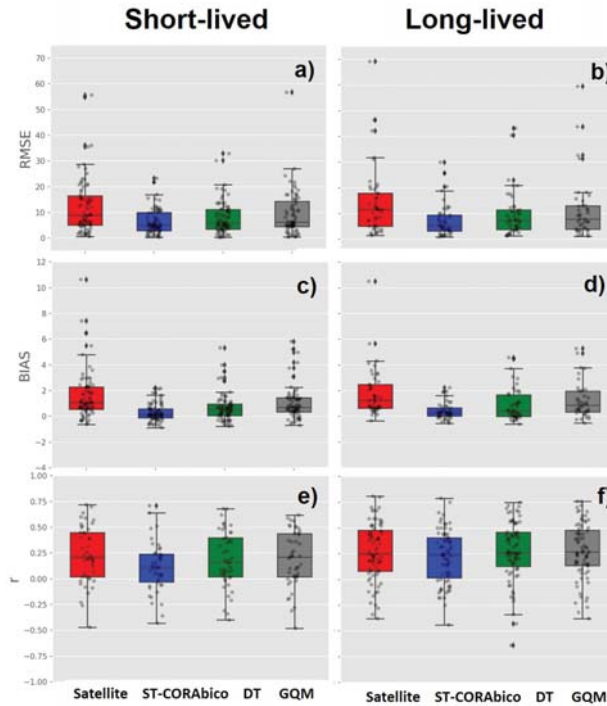
Figure 10. Comparison between the satellite GPM-IMERG (red), ST-CORAbico (blue), Distribution Transformation (green) and Gamma Quantile Mapping (grey) error dispersion during monsoon seasons 2014–2017. (**a**,**b**) RMSE; (**c**,**d**) bias; and, (**e**,**f**) correlation coefficient.

5. Discussion

ST-CORAbico is a spatiotemporal object-based bias correction method that was designed to reduce the displacement and volume systematic errors of storm events detected by SPP. In comparison to spatial object-based bias correction methods e.g., [37,38], the inclusion of the temporal component of the storm event reduced additional error effects due to timing and orientation, improving the efficiency of the bias correction.

This research incorporated a multivariate kernel distribution algorithm into ST-CORA to segment the storm event using the four dimensions of the storm event. In comparison to binary segmentation in the previous version, ST-CORA with KDE segmentation was able to delineate intense storm events by removing unreal storm configurations as a consequence of false merging and false separation of storms due to the multidimensional connected labelling component algorithm. Based on the analysis of KDE threshold delineation and the connected intensity, we found that storm events that were segmented by

the 25th percentile of the distribution showed a good result for segmenting intense storms with strong connection. However, further improvement is required.

The implementation of ST-CORAbico described the individual error correction due to displacement and volume. Results in the Lower Mekong basin indicated that volume errors were the main error correction, primarily resulting from the high overestimation of GPM-IMERG. These results agreed with multiple findings regarding hourly GPM-IMERG in monsoonal areas [70,71]. Overall, volume and displacement errors effectively contributed to the reduction of bias and RMSE, demonstrating the importance of reducing both of these systematic errors in satellite correction.

We acknowledge certain limitations of the study. Firstly, the uncertainty arising from the spatial interpolation method that was used for rain gauge values was not fully addressed in this research. Volume and especially displacement corrections in ST-CORAbico can be affected by the type of interpolation methods used to represent the spatiotemporal distribution of the observed storm. A dense rain gauge network can reduce the level of uncertainty; however, it is important to evaluate the impact of the type of interpolation method on the performance of ST-CORAbico, as mentioned above. Another limitation arises from the sensitivity of *IoU* percentage to match observed and estimated storm events. Higher levels do not always correspond to similar events, which affects the bias correction. This process required an in-depth sensitivity analysis of *IoU* in order to reduce the automatic storm matching. Additional analysis is required in order to identify why there is a strong correlation between observed and predicted storms in a spatiotemporal environment. In this study, we validated the performance of ST-CORAbico by comparing its performance against two widely used probabilistic methods. However, error metrics were calculated using the observed values, as there is no independent validation dataset available. Further implementations should consider an independent dataset to validate the error correction of the ST-CORAbico method.

This study was conducted in collaboration with the SERVIR-Mekong project and the Mekong River Commissions (MRC). SERVIR-Mekong is harnessing space and geospatial technologies to help decision-makers and key civil society groups to integrate geospatial information into their decision-making, planning, and communication. The application of this methodology can be used for various scientific purposes, including flood risk and water management. More specifically, the methodology enhances the input rainfall data, which are a crucial component of flood and drought early warning systems, landslide monitoring, as well as other water-related decision support systems. Future work will include the integration of machine learning technologies for near real-time bias correction of rainfall data when field data are scarce. In this regard, machine learning models will be trained and optimised using legacy field data and deployed on a near real-time basis.

6. Conclusions

We proposed a new spatiotemporal bias correction method for storm prediction detected by satellites. The method, called Spatiotemporal Contiguous Object-based Rainfall Analysis for bias correction (ST-CORAbico), analyses the main spatiotemporal characteristics of the observed and estimated storm events to correct systematic error sources due to displacement and volume. This methodology has two main elements: storm analysis for the segmentation and classification of storm event; and bias correction for correcting error due displacement and volume. In the storm analysis, we applied the ST-CORA method with a multivariate kernel segmentation in order to identify the spatiotemporal structure of the storm event. This method was applied over the Lower Mekong basin in Thailand to correct the GPM-IMERG Early version during the monsoon seasons from 2014 to 2017. The performance of ST-CORAbico was evaluated against the Distribution Transformation and the Gamma Quantile Mapping methods based on the reduction of RMSE, bias, and correlation coefficient. The results were divided by classifying the storm events into short- and long-lived storm events while using the k-means cluster analysis method.

We classified 68 storms (56%) as short-lived storms and 52 (44%) as long-lived events. The results of both storm event types showed that ST-CORAbico reduced the RMSE and bias of GPM-IMERG.

Volume correction was the major error source due to the overestimation present in GPM-IMERG. Location error was most important in the reduction of the bias. ST-CORAbico displayed a marginal impact on the spatial structure of the satellite-derived rainfall, showing the original structure of the rainfall data.

The comparison of ST-CORAbico with the Distribution Transformation and the Gamma Quantile Mapping methods showed that ST-CORAbico had the lowest RMSE as well as the lowest bias in both short and long-lived events. In terms of the correlation coefficient, ST-CORAbico and DT had a lower impact on the correlation coefficient, especially for short-lived events.

ST-CORAbico improves the accuracy of satellite-derived near real-time information on storm events. It can be used in various flood monitoring and water management applications. Our future studies will also incorporate machine learning methods and related technologies in order to correct storm events in real-time, in situations where field observation data are scarce.

References

1. Bui, Y.T.; Orange, D.; Visser, S.; Hoanh, C.T.; Laissus, M.; Poortinga, A.; Tran, D.T.; Stroosnijder, L. Lumped surface and sub-surface runoff for erosion modeling within a small hilly watershed in northern Vietnam. *Hydrol. Process.* **2014**, *28*, 2961–2974. [CrossRef]

2. Markert, K.N.; Schmidt, C.M.; Griffin, R.E.; Flores, A.I.; Poortinga, A.; Saah, D.S.; Muench, R.E.; Clinton, N.E.; Chishtie, F.; Kityuttachai, K.; et al. Historical and operational monitoring of surface sediments in the lower mekong basin using Landsat and Google Earth Engine cloud computing. *Remote Sens.* **2018**, *10*, 909. [CrossRef]

3. Poortinga, A.; Bastiaanssen, W.; Simons, G.; Saah, D.; Senay, G.; Fenn, M.; Bean, B.; Kadyszewski, J. A self-calibrating runoff and streamflow remote sensing model for ungauged basins using open-access earth observation data. *Remote Sens.* **2017**, *9*, 86. [CrossRef]

4. Price, K.; Purucker, S.T.; Kraemer, S.R.; Babendreier, J.E.; Knightes, C.D. Comparison of radar and gauge precipitation data in watershed models across varying spatial and temporal scales. *Hydrol. Process.* **2014**, *28*, 3505–3520. [CrossRef]

5. AghaKouchak, A.; Nakhjiri, N. A near real-time satellite-based global drought climate data record. *Environ. Res. Lett.* **2012**, *7*, 044037. [CrossRef]

6. Azarderakhsh, M.; Rossow, W.B.; Papa, F.; Norouzi, H.; Khanbilvardi, R. Diagnosing water variations within the Amazon basin using satellite data. *J. Geophys. Res. Atmos.* **2011**, *116*. [CrossRef]

7. Pan, M.; Li, H.; Wood, E. Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resour. Res.* **2010**, *46*. [CrossRef]

8. Tesfagiorgis, K.; Mahani, S.; Krakauer, N.; Khanbilvardi, R. Bias correction of satellite rainfall estimates using a radar-gauge product—A case study in Oklahoma (USA). *Hydrol. Earth Syst. Sci.* **2011** *15*, 2631–2647. [CrossRef]

9. Griffith, C.G.; Woodley, W.L.; Grube, P.G.; Martin, D.W.; Stout, J.; Sikdar, D.N. Rain estimation from geosynchronous satellite imagery—Visible and infrared studies. *Mon. Weather Rev.* **1978**, *106*, 1153–1171. [CrossRef]

10. Vicente, G.A.; Scofield, R.A.; Menzel, W.P. The operational GOES infrared rainfall estimation technique. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 1883–1898. [CrossRef]

11. Marzano, F.S.; Palmacci, M.; Cimini, D.; Giuliani, G.; Turk, F.J. Multivariate statistical integration of satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1018–1032. [CrossRef]

12. Tapiador, F. A physically based satellite rainfall estimation method using fluid dynamics modelling. *Int. J. Remote Sens.* **2008**, *29*, 5851–5862. [CrossRef]

13. Sorooshian, S.; Hsu, K.L.; Gao, X.; Gupta, H.V.; Imam, B.; Braithwaite, D. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2035–2046. [CrossRef]

14. 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]

15. 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]

16. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Nelkin, E.J. The TRMM multi-satellite precipitation analysis (TMPA). In *Satellite Rainfall Applications for Surface Hydrology*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 3–22.

17. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. *NASA/GSFC Code* **2015**, *612*, 2019.

18. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K.L. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.* **2018**, *56*, 79–107. [CrossRef]

19. Thiemig, V.; Rojas, R.; Zambrano-Bigiarini, M.; Levizzani, V.; De Roo, A. Validation of satellite-based precipitation products over sparsely gauged African river basins. *J. Hydrometeorol.* **2012**, *13*, 1760–1783. [CrossRef]

20. 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. *Remote Sens.* **2015**, *7*, 7181–7211. [CrossRef]

21. Kimani, M.W.; Hoedjes, J.C.B.; Su, Z. An Assessment of satellite-derived rainfall products relative to ground observations over East Africa. *Remote Sens.* **2017**, *9*, 430. [CrossRef]

22. Vila, D.A.; De Goncalves, L.G.G.; Toll, D.L.; Rozante, J.R. Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. *J. Hydrometeorol.* **2009**, *10*, 533–543. [CrossRef]

23. Habib, E.; Haile, A.T.; Sazib, N.; Zhang, Y.; Rientjes, T. Effect of bias correction of satellite-rainfall estimates on runoff simulations at the source of the Upper Blue Nile. *Remote Sens.* **2014**, *6*, 6688–6708. [CrossRef]

24. Ebert, E.; McBride, J. Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.* **2000**, *239*, 179–202. [CrossRef]

25. Davis, C.A.; Brown, B.G.; Bullock, R.; Halley-Gotway, J. The method for object-based diagnostic evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring Program. *Weather Forecast.* **2009**, *24*, 1252–1267. [CrossRef]

26. Wernli, H.; Paulat, M.; Hagen, M.; Frei, C. SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Weather Rev.* **2008**, *136*, 4470–4487. [CrossRef]

27. Roberts, N. An investigation of the ability of a storm scale configuration of the Met Office NWP model to predict floodproducing rainfall. *UK Met Off. Tech. Rep.* **2005**, *455*, 80.

28. Marsigli, C.; Montani, A.; Paccagnella, T. Verification of the COSMOLEPS new suite in terms of precipitation distribution. *COSMO Newsl.* **2006**, *6*, 134–141.

29. Casati, B.; Ross, G.; Stephenson, D. A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl. A J. Forecast. Pract. Appl. Train. Tech. Model.* **2004**, *11*, 141–154. [CrossRef]

30. Casati, B. New developments of the intensity-scale technique within the Spatial Verification Methods Intercomparison Project. *Weather Forecast.* **2010**, *25*, 113–143. [CrossRef]

31. Mittermaier, M.P. Using an intensity-scale technique to assess the added benefit of high-resolution model precipitation forecasts. *Atmos. Sci. Lett.* **2006**, *7*, 36–42. [CrossRef]

32. Keil, C.; Craig, G.C. A displacement-based error measure applied in a regional ensemble forecasting system. *Mon. Weather Rev.* **2007**, *135*, 3248–3259. [CrossRef]

33. Keil, C.; Craig, G.C. A displacement and amplitude score employing an optical flow technique. *Weather Forecast.* **2009**, *24*, 1297–1308. [CrossRef]

34. Skok, G.; Tribbia, J.; Rakovec, J.; Brown, B. Object-based analysis of satellite-derived precipitation systems over the low-and midlatitude Pacific Ocean. *Mon. Weather Rev.* **2009**, *137*, 3196–3218. [CrossRef]

35. Li, J.; Hsu, K.; AghaKouchak, A.; Sorooshian, S. An object-based approach for verification of precipitation estimation. *Int. J. Remote Sens.* **2015**, *36*, 513–529. [CrossRef]

36. Li, J.; Hsu, K.L.; AghaKouchak, A.; Sorooshian, S. Object-based assessment of satellite precipitation products. *Remote Sens.* **2016**, *8*, 547. [CrossRef]

37. Demaria, E.; Rodriguez, D.; Ebert, E.; Salio, P.; Su, F.; Valdes, J.B. Evaluation of mesoscale convective systems in South America using multiple satellite products and an object-based approach. *J. Geophys. Res. Atmos.* **2011**, *116*. [CrossRef]

38. Le Coz, C.; Heemink, A.; Verlaan, M.; Veldhuis, M.C.; van de Giesen, N. Correcting position error in precipitation data using image morphing. *Remote Sens.* **2019**, *11*, 2557. [CrossRef]

39. Clark, A.J.; Kain, J.S.; Marsh, P.T.; Correia, J., Jr.; Xue, M.; Kong, F. Forecasting tornado pathlengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts. *Weather Forecast.* **2012**, *27*, 1090–1113. [CrossRef]

40. Clark, A.J.; Bullock, R.G.; Jensen, T.L.; Xue, M.; Kong, F. Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Weather Forecast.* **2014**, *29*, 517–542. [CrossRef]

41. Development and implementation of MODE time domain object-based verification In Proceedings of the 24th Conference Weather and Forecasting, Seattle, WA, USA, 24–27 January 2011.

42. Laverde-Barajas, M.; Corzo Perez, G.; Chishtie, F.; Poortinga, A.; Uijlenhoet, R.; Solomatine, D. Decomposing satellite-based rainfall errors in flood estimation: Hydrological responses using a spatiotemporal object-based verification method. *J. Hydrol.* **2020**, 125554. [CrossRef]

43. Laverde-Barajas, M.; Corzo, G.; Bhattacharya, B.; Uijlenhoet, R.; Solomatine, D.P. Spatiotemporal Analysis of Extreme Rainfall Events Using an Object-Based Approach. In *Spatiotemporal Analysis of Extreme Hydrological Events*; Corzo, G., Varouchakis, E.A., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 95–112. [CrossRef]

44. Delgado, J.; Merz, B.; Apel, H. A climate-flood link for the lower Mekong River. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 1533–1541. [CrossRef]

45. Yang, R.; Zhang, W.K.; Gui, S.; Tao, Y.; Cao, J. Rainy season precipitation variation in the Mekong River basin and its relationship to the Indian and East Asian summer monsoons. *Clim. Dyn.* **2019**, *52*, 5691–5708. [CrossRef]

46. Schneider, U.; Fuchs, T.; Meyer-Christoffer, A.; Rudolf, B. *Global Precipitation Analysis Products of the GPCC*; Global Precipitation Climatology Centre (GPCC), DWD, Internet Publication; Deutscher Wetterdienst: Offenbach, Germany, 2008; Volume 112.

47. Golden Software, I. *Surfer 12 Users' Guide: Powerful Contouring, Gridding, and Surface Mapping*, 12th ed.; Golden Software, Inc.: Golden, CO, USA, 2014.

48. Ly, S.; Sohier, C.; Charles, C.; Degré, A. Effect of raingage density, position and interpolation on rainfall-discharge modelling. In *EGU2012 Geophysical Research Abstracts*; European Geophysical Society: Katlenburg-Lindau, Germany, 2012; p. 2592.

49. Li, J.; Heap, A.D. Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Model. Softw.* **2014**, *53*, 173–189. [CrossRef]

50. Chen, Y.C.; Wei, C.; Yeh, H.C. Rainfall network design using kriging and entropy. *Hydrol. Process. Int. J.* **2008**, *22*, 340–346. [CrossRef]

51. Guttman, A. R-Trees: A Dynamic Index Structure for Spatial Searching. In Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data, Association for Computing Machinery, New York, NY, USA, 18–21 June 1984; pp. 47–57. Available online: http://www-db.deis.unibo.it/courses/SI-LS/papers/Gut84.pdf (accessed on 30 September 2020).

52. Ebert, E.E.; Gallus, W.A., Jr. Toward better understanding of the contiguous rain area (CRA) method for spatial forecast verification. *Weather Forecast.* **2009**, *24*, 1401–1415. [CrossRef]

53. Bethel, E.W.; Childs, H.; Hansen, C. *High Performance Visualization: Enabling Extreme-Scale Scientific Insight*; CRC Press: Boca Raton, FL, USA, 2012.

54. Chen, T.B.; Lu, H.H.S.; Lee, Y.S.; Lan, H.J. Segmentation of cDNA microarray images by kernel density estimation. *J. Biomed. Inf.* **2008**, *41*, 1021–1027. [CrossRef]

55. Pereira, O.; Torres, E.; Garcés, Y.; Rodríguez, R. Edge detection based on kernel density estimation. In Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), The Steering Committee of The World Congress in Computer Science, Computer, Engineering and Applied Computing WorldComp, Las Vegas, NV, USA, 17–20 July 2017; pp. 123–128.

56. Wang, L.; Lu, J.; Li, X.; Huan, Z.; Liang, J.; Chen, S. Learning arbitrary-shape object detector from bounding-box annotation by searching region-graph. *Pattern Recognit. Lett.* **2017**, *87*, 171–176. [CrossRef]

57. Zivkovic, Z.; Van Der Heijden, F. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognit. Lett.* **2006**, *27*, 773–780. [CrossRef]

58. Berjón, D.; Cuevas, C.; Morán, F.; García, N. Real-time nonparametric background subtraction with tracking-based foreground update. *Pattern Recognit.* **2018**, *74*, 156–170. [CrossRef]

59. Hyndman, R.L.; Zhang, X.; King, M.L. Bandwidth selection for multivariate kernel density estimation using MCMC. In *Econometric Society 2004 Australasian Meetings*; Econometric Society: Melbourne, Australia, 7–9 July 2004.

60. Zhang, X.; King, M.L.; Hyndman, R.J. A Bayesian approach to bandwidth selection for multivariate kernel density estimation. *Comput. Stat. Data Anal.* **2006**, *50*, 3009–3031. [CrossRef]

61. Henderson, D.J.; Parmeter, C.F. Normal reference bandwidths for the general order, multivariate kernel density derivative estimator. *Stat. Probab. Lett.* **2012**, *82*, 2198–2205. [CrossRef]

62. Grams, J.S.; Gallus, W.A., Jr.; Koch, S.E.; Wharton, L.S.; Loughe, A.; Ebert, E.E. The use of a modified Ebert–McBride technique to evaluate mesoscale model QPF as a function of convective system morphology during IHOP 2002. *Weather Forecast.* **2006**, *21*, 288–306. [CrossRef]

63. Steiner, M.; Houze, R.A., Jr.; Yuter, S.E. Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *J. Appl. Meteorol.* **1995**, *34*, 1978–2007. [CrossRef]

64. Molini, L.; Parodi, A.; Siccardi, F. Dealing with uncertainty: An analysis of the severe weather events over Italy in 2006. *Nat. Hazards Earth Syst. Sci* **2009**, *9*, 1775–1786. [CrossRef]

65. Molini, L.; Parodi, A.; Rebora, N.; Craig, G. Classifying severe rainfall events over Italy by hydrometeorological and dynamical criteria. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 148–154. [CrossRef]

66. Johnson, A.E.; Hebert, M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 433–449. [CrossRef]

67. Themeßl, M.J.; Gobiet, A.; Heinrich, G. Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Clim. Chang.* **2012**, *112*, 449–468. [CrossRef]

68. Brown, B.G.; Bullock, Y.R.; Davis, C.A.; Gotway, J.H.; Chapman, M.B.; Takacs, A.; Gillel, E.; Manning, K.; Mahoney, J.L. New verification approaches for convective weather forecasts. In Proceedings of the 11th Conference on Aviation, Range, and Aerospace, Hyannis, MA, USA, 4–8 October 2004; pp. 3–8.

69. Piani, C.; Haerter, J.; Coppola, E. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.* **2010**, *99*, 187–192. [CrossRef]

70. Oliveira, R.; Maggioni, V.; Vila, D.; Morales, C. Characteristics and diurnal cycle of GPM rainfall estimates over the central Amazon region. *Remote Sens.* **2016**, *8*, 544. [CrossRef]

71. 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]

Comparison of GPM IMERG and TRMM 3B43 Products over Cyprus

Adrianos Retalis [1,*], Dimitris Katsanos [1], Filippos Tymvios [2,3] and Silas Michaelides [3]

[1] Institute for Environmental Research & Sustainable Development, National Observatory of Athens, GR15236 Athens, Greece; katsanos@noa.gr
[2] Department of Meteorology, 1086 Nicosia, Cyprus; ftymvios@dom.moa.gov.cy
[3] Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, 2121 Aglantzia, Cyprus; s.michaelides@cyi.ac.cy
*   Correspondence: adrianr@noa.gr; Tel.: +30-210-8109201

Abstract: Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) high-resolution product and Tropical Rainfall Measuring Mission (TRMM) 3B43 product are validated against rain gauges over the island of Cyprus for the period from April 2014 to June 2018. The comparison performed is twofold: firstly, the Satellite Precipitation (SP) estimates are compared with the gauge stations' records on a monthly basis and, secondly, on an annual basis. The validation is based on ground data from a dense and well-maintained network of rain gauges, available in high temporal (hourly) resolution. The results show high correlation coefficient values, on average reaching 0.92 and 0.91 for monthly 3B43 and IMERG estimates, respectively, although both IMERG and TRMM tend to underestimate precipitation (Bias values of −1.6 and −3.0, respectively), especially during the rainy season. On an annual basis, both SP estimates are underestimating precipitation, although IMERG estimates records (R = 0.82) are slightly closer to that of the corresponding gauge station records than those of 3B43 (R = 0.81). Finally, the influence of elevation of both SP estimates was considered by grouping rain gauge stations in three categories, with respect to their elevation. Results indicated that both SP estimates underestimate precipitation with increasing elevation and overestimate it at lower elevations.

Keywords: GPM; IMERG; TRMM; precipitation; Cyprus

1. Introduction

Satellite observations have been widely used during recent decades for several meteorological, hydrological and climatological applications incorporating precipitation data worldwide [1–6]. In order to fill in where ground observations are absent or sparse, satellite estimations have been evolving using sophisticated algorithms that can identify rainfall, snow and/or other hydrometeors [7–11]. However, although Satellite Precipitation (SP) products are able to, overall, capture the variability and magnitude of rainfall, still they cannot accurately estimate the localized rainfall variations. Thus, validation of satellite precipitation products is often needed against ground-based measurements.

The Tropical Rainfall Measuring Mission (TRMM) platform placed in orbit during the 1997–2015 period, provided reliable data of high spatial (≈25 km) and temporal resolution (3 h), at a geographical coverage between 50° N and 50° S [12–17]. TRMM's successor, namely, the Global Precipitation Mission (GPM) has been in orbit since 2014, giving estimates at even higher resolutions (≈10 km; 30 min) and geographical coverage from 60° S to 60° N, making it available for a variety of applications, including the assimilation of GPM data in numerical weather prediction models to improve model forecasting skill [15,16], the monitoring of severe weather events [13,15–17], hydrological hazards [18,19], etc.

Several studies attempted to demonstrate the accuracy of TRMM and GPM IMERG (Integrated Multi-satellitE Retrievals) estimates in various geographical areas. In their study over mainland China, Wu et al. [20] found that both SP products overestimate light rainfall. This is attributed to the fact that hydrometeors detected by infrared and microwave sensors as well as precipitation radars may partially or even totally evaporate before they are registered by the rain gauges. Furthermore, these authors found an underestimation of moderate and heavy rainfall by both products. A slightly better performance by GPM-IMERG, according to these authors, is attributed to the satellite overpasses and sensor capabilities.

In a similar study over Pakistan, Anjum et al. [21] found a slight dominance of IMERG; however, both products correlate well with the in situ measurements at a monthly scale, adequately following the temporal pattern. Again, underestimation of moderate and heavy rainfall and overestimation of light events was reported.

In their study regarding the area of Singapore, Tan and Duan [22] presented similar results, showing good correlation on a monthly scale, with rain gauges for both products and moderate correlation for daily values. The authors underlined that the better performance of IMERG was not that notable and that the main advantage of the new product was mostly its finer resolution.

In a study over the Tibetan plateau (Hexi region), Wang et al. [23] found a better correlation for IMERG, ascribed mostly to the ability to detect better moderate and heavy rainfall; however, they concluded that the improvement was not significant.

In their study (China, 2015–2017), Chen et al. [24] evaluated the performance of IMERG (v5) and TRMM 3B42 (v7) and found that, at monthly and annual scale, both datasets were highly correlated with rain gauge observations. Considering daily values, satellite estimates overestimate precipitation for intensities within the range 0 to 25 mm/day and underestimate precipitation for light and heavy intensities. Considering various statistical scores, they found that IMERG, in general, performed better in detecting the observed precipitation.

In a similar study (China, March 2014 to February 2017), Wei et al. [25] found severe underestimation with high negative relative biases for both IMERG (v5) and TRMM products. However, IMERG product performed better than TRMM 3B42 in the detection of precipitation events in terms of specific statistical scores (i.e., probability of detection), over China and across most of the sub-regions.

Sunilkumar et al. [26] evaluated the GPM-IMERG (v5) final precipitation product against a ground-based gridded data set over Japan, Nepal and the Philippines for two years (2014–2015). Their results showed generally good performance (in terms of statistical scores, like correlation, mean bias, root mean square error) of GPM-IMERG over three regions, although an underestimation was noticed during heavy rainfall events. They also noticed that GPM-IMERG estimates improved its capability in terms of detecting light and heavy precipitation events, although their performance was found to be seasonally dependent.

A few studies with evaluation of satellite precipitation products over Cyprus are reported in the literature. Retalis et al. [27] performed an analysis of precipitation data from satellite data TRMM 3B43 (versions 7 and 7A) over Cyprus and compared them with the corresponding gauge observations and E-OBS gridded data (i.e., a European daily high-resolution gridded dataset of surface temperature and precipitation to be used for validation of Regional Climate Models and for climate change studies) for a 15-year period (1998–2012). They concluded that correlation between TRMM and E-OBS was higher in summertime (≈0.97), but significantly lower in the winter period (≈0.60–0.74). It was noticed that the annual correlation tends to decrease considerably with time. They also found that the coefficient of determination between TRMM, E-OBS estimates and gauge data were relatively high (0.929 and 0.932, respectively); however, the variations noticed were attributed to the elevation differences.

A study for a 30-year period (1981–2010) for the precipitation database Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) in Cyprus was presented by Katsanos et al. [28]. The CHIRPS database was evaluated against gauge stations data. Results showed good correlation between monthly CHIRPS values and recorded precipitation with the correlation coefficients found

to be around 0.85 and January the month with the highest correlation. The corresponding values for the annual mean ranged between 0.70 and 0.74, with the mountainous stations showing a slightly higher correlation.

In a later study, Katsanos et al. [29] examined the performance of several climatic indices for the CHIRPS precipitation dataset and rain gauges records on high spatial (0.05°) and temporal (daily) resolution for a period of 30 years (1981–2010). Results indicates quite a promising performance regarding indices related to daily precipitation thresholds, resulting in high correlation scores. However, for indices referring to number of days, results showed medium or no correlation, probably due to the criteria used for the identification of a wet (rainy) day on the CHIRPS dataset.

Furthermore, Retalis et al. [30], in their study on the accuracy of the GPM IMERG estimates over Cyprus (April 2014 to February 2017), concluded that, overall, a very good agreement (based on the statistical analysis) between monthly IMERG estimates and gauge data was established (coefficient of determination r^2 value $\approx$ 0.93), presenting a tendency of IMERG for underestimation when higher elevation (>1000 m) was considered. They also examined the daily dependency of IMERG estimates and gauge data, considering a series of extreme precipitation events, and they concluded that this is case dependency, while elevation does not have an apparent effect.

The objective of this study is to evaluate statistically the performance and improvement of the GPM IMERG product compared to TRMM 3B43V7 estimates, thus exploring, the continuity and uniformity between IMERG and TRMM-era data sets over Cyprus so that they can be used in climate studies as a combined and consistent dataset. The present research is a continuation and extension of previous studies by the same authors.

The current research aims at comparing the two products, namely, GPM IMERG and TRMM 3B43, in order to determine and highlight possible differences, advantages and disadvantages of each one of them, based on the performance of several statistical skill scores, along with cross-evaluation against the dense rain gauge dataset over Cyprus during the period from April 2014 to June 2018.

2. Study Area

Located in the north-eastern corner of the Mediterranean Sea, the island of Cyprus has a typical eastern-Mediterranean climate. The major characteristic of this type of climate can be concisely described by a bimodal seasonality with alternating relatively short wet winters and prolonged dry summers. As can be seen from the geomorphological map in Figure 1, the island is transversed by two mountain ranges: the high Troodos massif in the southwest with the highest peak, Olympus at 1951 m, and the elongated east-west oriented narrow Pentadaktylos range, rising to 900 m which borders the northern coast from east to west. Between the two mountain ranges, lies the central Mesaoria plain. Narrow, relatively flat strips of land surround the island along its coast.
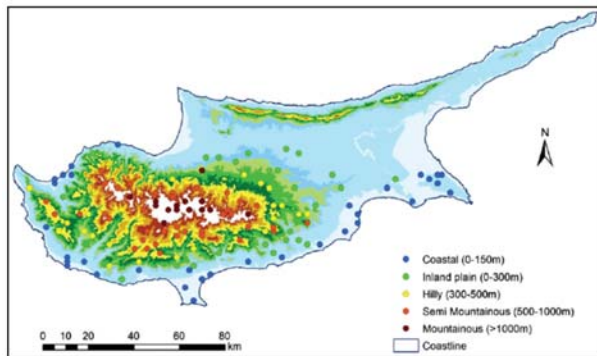
Figure 1. Geomorphology map of Cyprus highlighting the distribution of the dense gauge station network grouped according to elevation.

Most of the winter dynamic systems which affect Cyprus originate from the southwest to west [31,32]; hence, the highest average annual precipitation values are recorded on the southern side of the highest peaks of the Troodos mountain and the lowest over the rain-shadowed areas north of Troodos and at coastal stations on the east part of the island [33].

3. Data and Methodology

3.1. In-Situ Rain Data

The in-situ meteorological stations of the very reliable [30] dense and well-distributed network operated by the Cyprus Department of Meteorology were used for ground validation. Daily and hourly complete data records from 136 rain gauges were used for the study period (see Figure 1). These gauges are distributed in such a way so as to cover the whole study area, including not only coastal, urban and agricultural areas, but also the hilly and mountainous areas. It is worth noting that data underwent quality control prior to the present analysis.

The average annual total precipitation increases up the southwestern windward slopes from 450 mm to nearly 1100 mm at the top of the central massif. On the leeward slopes amounts decrease steadily northwards and eastwards to between 300 and 350 mm in the central plain and the flat southeastern parts of the island [34].

The annual average rainfall, covering the period 1951–1980, is 480 mm, exhibiting a decreasing trend in the last 30 years [34]. Furthermore, rainfall in the warmer months contributes little or nothing to the annual precipitation amounts.

It should be noted at this point that data from four meteorological stations are provided by the Cyprus Department of Meteorology to be incorporated in the TRMM and GPM estimates. These four stations are not representative of the country as a whole. Furthermore, they barely make up 3% of the data used herein, hence, the are not expected to have an impact on the objectivity of the study.

3.2. IMERG Data

Global Precipitation Measurement (GPM) mission was launched on 27 February 2014, as a successor of the Tropical Rainfall Measuring Mission (TRMM). GPM constellation incorporates passive microwave (PMW) and infrared (IR) satellites, providing global precipitation measurements within the range 60° N–60° S and better temporal and spatial analysis (see [25,35,36]).

GPM consists of one Core Observatory and approximately 10 constellation satellites. The Core Observatory carries a Ku/Ka-band dual-frequency precipitation radar and a multi-channel GPM microwave imager, extending the measurement range of TRMM instruments. GPM provides three levels of precipitation-related products. The level-3 products are produced with the IMERG (Integrated Multi-satellitE Retrievals for GPM) algorithm, which intercalibrates and merges precipitation estimates from all constellation microwave sensors, microwave-calibrated infrared satellite estimates, and monthly gauge precipitation data [37,38]. It is important here to comment on the PMW sensitivity of retrieved rainfall, since the launch of GPM, the overland rainfall retrieval algorithm, is transitioning from an inversion technique based on rainfall–brightness temperature scattering relationships to a Bayesian framework consistent with the over-ocean algorithm [39]. GPM IMERG precipitation estimates are available from 12 March 2014 to present. The GPM IMERG products offer a relatively fine spatial resolution of $0.1° \times 0.1°$ and high temporal resolution of 30 min, with a spatial coverage from 60° S to 60° N. The IMERG Final product [18] was chosen for our study and so, especially, was the IMERG (v05B) data for the period from April 2014 to June 2018.

3.3. TRMM Data

The Tropical Rainfall Measuring Mission (TRMM), launched by NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency) in 1997, provided precipitation estimates within the latitude 50° S to 50° N (see [40,41]).

The TRMM satellite carried several instruments to detect precipitation, including the Visible Infrared Radiometer (VIRS), TRMM Microwave Imager (TMI), Cloud and Earth Radiant Energy Sensor (CERES), Lightning Imaging Sensor (LIS) and the first spaceborne precipitation radar (PR). Several precipitation retrieval algorithms have been developed based on observations from the sensors on board the TRMM satellite such as the TMPA (TRMM Multi-satellite Precipitation Analysis). The TMPA algorithm combines observations from satellite-based microwave and infrared sensors and ground rainfall gauge analyses, and produces 3-hourly rainfall estimates at a spatial resolution of $0.25° \times 0.25°$ with a quasi-global coverage (50° N-S) [42]. In 2015, the TRMM mission came to an end, with the instruments turned off and the spacecraft re-entering the Earth's atmosphere. However, the multi-satellite TMPA products continue to be produced using input data from other satellites in the constellation. Indeed, the TMPA algorithms are still being run using other calibrators to produce data in parallel with GPM IMERG [43].

The Level 3 TRMM 3B43 data, also called TMPA product, were chosen for our analysis [42]. In particular, TRMM 3B43 (v7) data for the period from April 2014 to June 2018 were used.

3.4. Methods

In order to perform the evaluation of IMERG and 3B43V7 products relative to the reference rain gauges data, several indices including Pearson Correlation Coefficient (R), mean error (Bias), relative Bias (rBias), Root Mean Square Error (RMSE) and mean absolute error (MAE) were computed (see Table 1). Pearson correlation coefficient (R) is a dimensionless statistical index used to assess the linear correlation between the reference ground-based data and the satellite precipitation estimates. Mean error (Bias) represents the systematic error of satellite precipitation estimates, a measure of the overestimation or underestimation of the gauge data. Relative Bias (rBias) estimates the relative difference (in percentage) between the two data sources (satellite estimates and rain gauges). RMSE quantifies the average error magnitude (mm/time) between the satellite estimates and the rain gauge data. Mean absolute error (MAE) reflects the magnitude and extent of the mean error of satellite precipitation estimates. For seasonal analysis, the year was divided into four seasons: winter (December to February); spring (March to May); summer (June to September); and autumn (October to November).

Table 1. Summary of statistical indices used to evaluate the satellite precipitation products (S$_i$: satellite estimates, Oi: observations).

	Unit	Equation
R (correlation coefficient)	-	$\dfrac{\sum_{i=1}^{n}(O_i-\overline{O})(S_i-\overline{S})}{\sqrt{\sum_{i=1}^{n}(O_i-\overline{O})^2}\sqrt{\sum_{i=1}^{n}(S_i-\overline{S})^2}}$
Mean Error (Bias)	mm	$\dfrac{\sum_{i=1}^{n}(S_i-O_i)}{n}$
Relative Bias	-	$\dfrac{\sum_{i=1}^{n}(S_i-O_i)}{\sum_{i=1}^{n}O_i}(100)$
RMSE	mm	$\sqrt{\dfrac{\sum_{i=1}^{n}(S_i-O_i)^2}{n}}$
Mean Absolute Error	mm	$\dfrac{\sum_{i=1}^{n}|S_i-O_i|}{n}$

Figure 2 shows the number of stations that are distributed within the grid cells of each Satellite Precipitation (SP) dataset. The number of grid cells for the study area was 19 for the TRMM 3B43 data and 61 for IMERG data, respectively. On the one hand, the TRMM 3B43 grids show a notable variation of the available number of gauge stations residing (e.g., from 1 to 23 stations per grid cell). It should be also noted that 42% of the available gauge stations (57 of 136) were located within only three cells, although these 3B43 cells are not the ones with the maximum correlation with the corresponding gauge

values. We notice that the distribution of the available gauge station within each IMERG grid cell was more balanced (1 to 5 stations per grid cell).
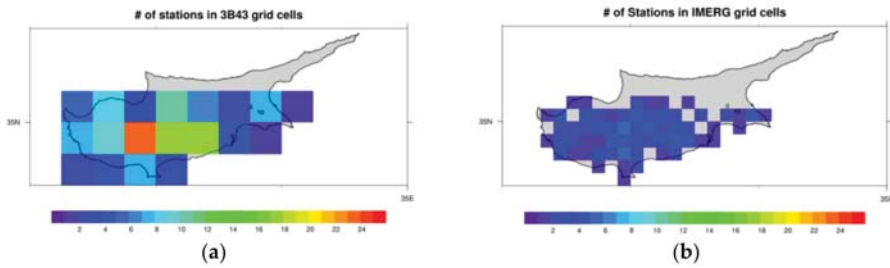
Figure 2. Distribution of rain gauge stations within Tropical Rainfall Measuring Mission (TRMM) 3B43 (**a**) and Integrated Multi-satellitE Retrievals (IMERG) (**b**) grids.

4. Results and Discussion

4.1. Monthly Validation

Mean monthly values for the study period were calculated for both the gauge stations (136 stations) and the two satellite precipitation products (mean values of all available corresponding grid cells within the study area). The results are illustrated in Figure 3. We notice that, overall, both IMERG and TRMM data follow very well the "climatology" of the stations, although with an underestimation during the rainy period, while IMERG is closer to the gauge values for almost the whole period of study.
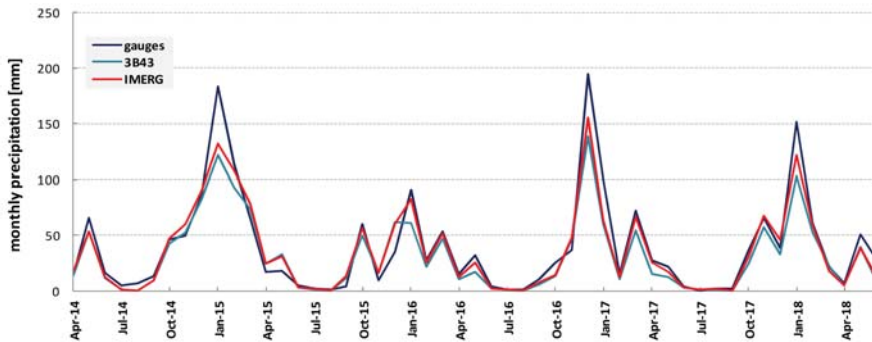
Figure 3. Average monthly values calculated for gauges (dark blue) TRMM 3B43 estimates (light blue) and IMERG estimates (red line) for the period April 2014 to June 2018.

In order to highlight the performance of satellite precipitation (SP) products, we estimate their average difference (of their monthly values) from the corresponding gauge stations data (see Figure 4). We notice that, generally, IMERG monthly values present lower divergence than TRMM.

Next, IMERG and TRMM monthly estimates were compared to the corresponding gauge data based on a grid-level approach. Thus, for each SP product, the mean gauge station value is calculated for the comparison if more than one station was located within each SP grid.
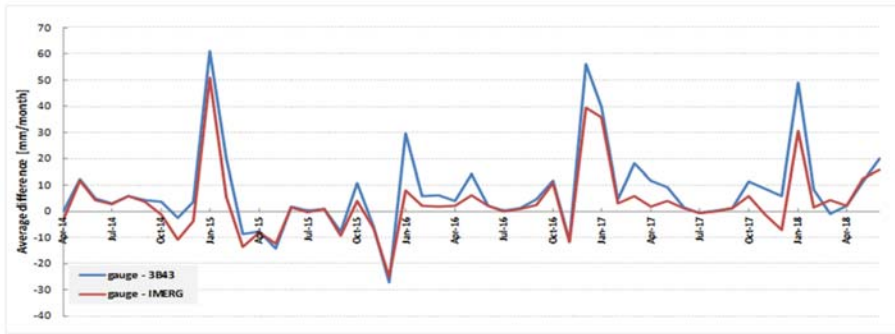
Figure 4. Average difference (monthly values) between gauge stations data from 3B43 (blue line) and IMERG (red line) estimates, respectively.

The estimated correlation coefficient values between the monthly values of the gauge stations and the SP data are presented in Figure 5. Correlation seems slightly better for TRMM 3B43 cells, since there is a lower variation (minimum 0.84–maximum 0.96) than that corresponding to the IMERG cells (minimum 0.78–maximum 0.96), with the average of all cells being 0.92 and 0.91, respectively.
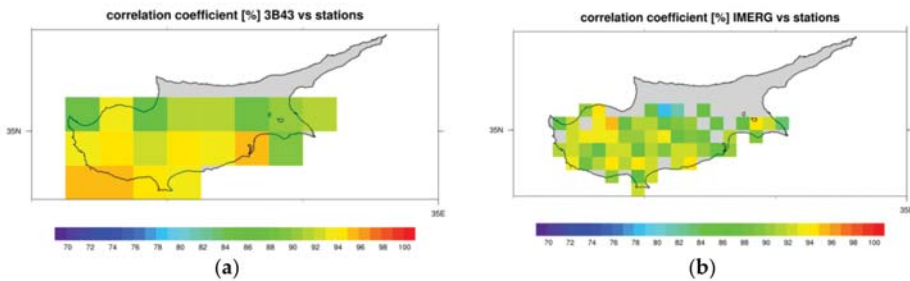
Figure 5. Correlation coefficient at grid cell level between the gauge stations and the TRMM (**a**) and IMERG (**b**) data.

Both IMERG and 3B43 products underestimate precipitation with Bias values of −1.6 mm/month and −3.0 mm/month, respectively. IMERG showed better performance than 3B43 in terms of rBias values (3.6 and −8.7, respectively), while presenting worst performance in terms of RMSE (23.8 mm/month and 20.0 mm/month, respectively) and MAE (15.1 mm/month and 12.9 mm/month, respectively) values, as presented in Table 2 (minimum and maximum respective values are provided in brackets).

Table 2. Performance evaluation metrics of monthly data from 3B43V7 and IMERG (the range of values is given in parentheses).

	TRMM 3B43	GPM IMERG
R (correlation coefficient)	0.92 (0.84–0.96)	0.91 (0.78–0.96)
Mean Error (Bias)	−3.0 (−21.5–7.7)	−1.6 (−31.8–13.0)
Relative Bias	−3.0 (−38.6–35.0)	3.6 (−45.6–58.7)
RMSE	20.0 (11.6–37.2)	23.8 (14.7–55.0)
Mean Absolute Error	12.9 (8.7–24.3)	15.1 (10.2–35.5)

4.2. Seasonal Validation

The results for the seasonal validation are summarized in Table 3. Both TRMM 3B43 and GPM IMERG estimates presented high R values (0.91 and 0.90, respectively) in winter, which suggests a good agreement with ground-based measurements on a seasonal scale. Similar results were found for autumn (0.84 and 0.83, respectively) and spring (0.81 and 0.80, respectively), while the poorest correlation (0.68 and 0.67, respectively) was established in summer.

Table 3. Performance evaluation metrics of seasonal data from 3B43V7 and IMERG (the range of values is given in parentheses).

	TRMM 3B43				GPM IMERG			
	MAM	**JJA**	**SON**	**DJF**	**MAM**	**JJA**	**SON**	**DJF**
R (correlation coefficient)	0.81	0.68	0.84	0.91	0.80	0.67	0.83	0.90
Mean Error (Bias)	0.9	−1.3	0.3	−12.6	0.8	−1.8	2.6	−8.3
Relative Bias	15.0	440.7	6.4	−9.9	17.5	299.8	18.4	−1.1
RMSE	15.2	6.3	13.9	33.4	19.1	8.7	16.7	38.6
Mean Absolute Error	12.2	3.8	11.0	25.6	14.6	4.9	12.7	29.2

The BIAS for TRMM 3B43 and IMERG ranged from −12.6 to 0.9 and from −8.3 to 2.6, respectively, in the four seasons. The 3B43 overestimated precipitation in spring (0.9) and autumn (0.3), while underestimation is noticed in summer (−1.3) and is rather significant in winter (−12.6). Similar is the pattern for IMERG, with overestimation in spring (0.8) and autumn (2.6), while underestimation is noticed in summer (−1.8) and is rather significant in winter (−8.3).

In terms of rBias, 3B43 presented larger values than that of IMERG in spring (15.0 and 17.5, respectively) and autumn (6.4 and 18.4, respectively), while IMERG showed better performance in winter (−1.1 and −9.9, respectively). Both products displayed their worst values in summer (440.7 for 3B43 and 299.8 for IMERG, respectively).

Precipitation products displayed a similar trend for RMSE and MAE with higher values in winter, spring and autumn and lower values in summer. For winter, spring and autumn, TRMM 3B43 had RMSE values of 33.4, 15.2 and 13.9, which were slightly lower than those of IMERG, which were 38.6, 19.1 and 16.7, respectively. In summer, 3B43 had lower RMSE values, 6.3, compared to those of IMERG, 8.7, respectively. Similar results occurred for MAE values between 3B43 and IMERG, with larger values noticed in winter (33.4 and 38.6, respectively) and lower in summer (3.8 and 4.9, respectively). These results may indicate that both SP products exhibit a similar error level, on a seasonal scale.

4.3. Annual Validation

Mean annual values for the study period (see Figure 6) were calculated for both the set of 136 ground stations and the satellite precipitation products (mean values of the available corresponding grid cells within the study area). It is found that, overall, both SP data exhibit an underestimation, although it is lower for IMERG, with the exception of 2015, when a slight overestimation by the IMERG product was noticed. It should be noted, however, that the available precipitation estimates for 2014 were limited to the period April to December, while for 2018 they were limited to January to June.
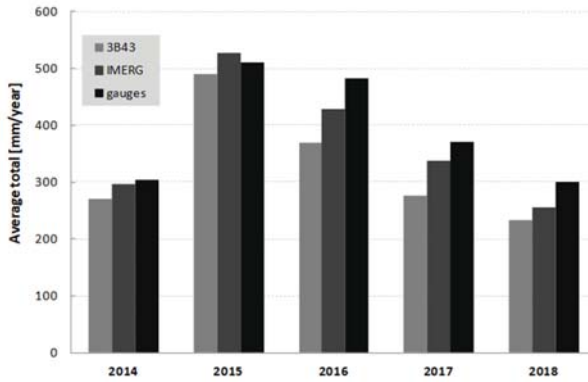
Figure 6. Mean annual values of gauge stations and Satellite Precipitation (SP) data.

Figure 7 displays the variation of overestimation (>100%) or underestimation (<100%) of annual precipitation values at each of the corresponding SP grid cell, between the SP estimated and calculated mean annual gauge station records. We notice that, more or less, the SP estimates have the same behavior regarding the overestimation/ underestimation of annual rainfall. Underestimation is more evident in the central area of Cyprus (greater area of Troodos mountain range), where higher precipitation records are generally noticed, highlighting the known limitations of satellite products regarding heavy rainfall, while the overestimation is noticed in the coastal or rather flat areas, where again the estimation of precipitation still remains a challenge due to the difficulty in distinguishing between rain and non-rain pixels over a complex background [44]. Furthermore, there are no significant differences between the years, since the cells that generally overestimate/underestimate rainfall have a similar performance regardless of the year.
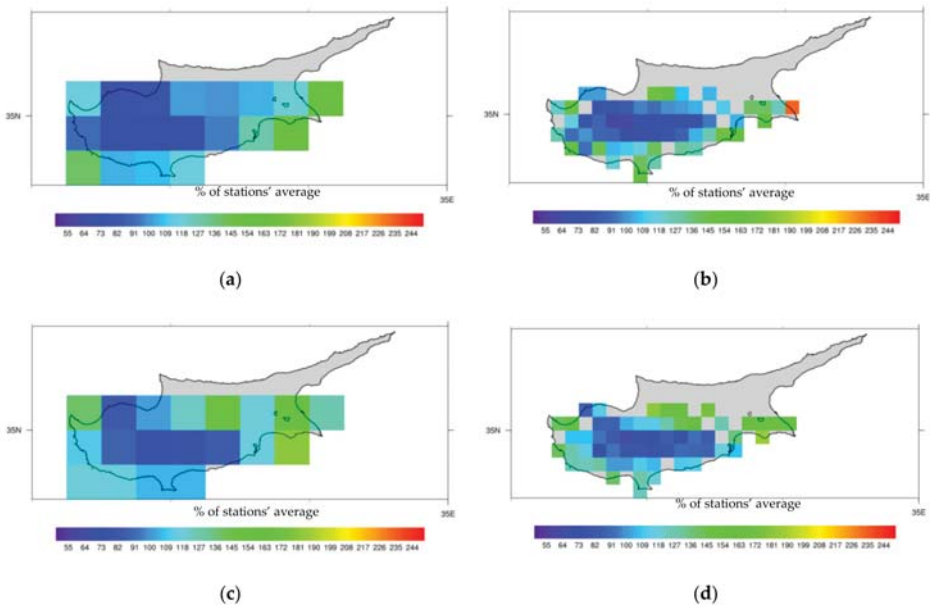
Figure 7. *Cont.*
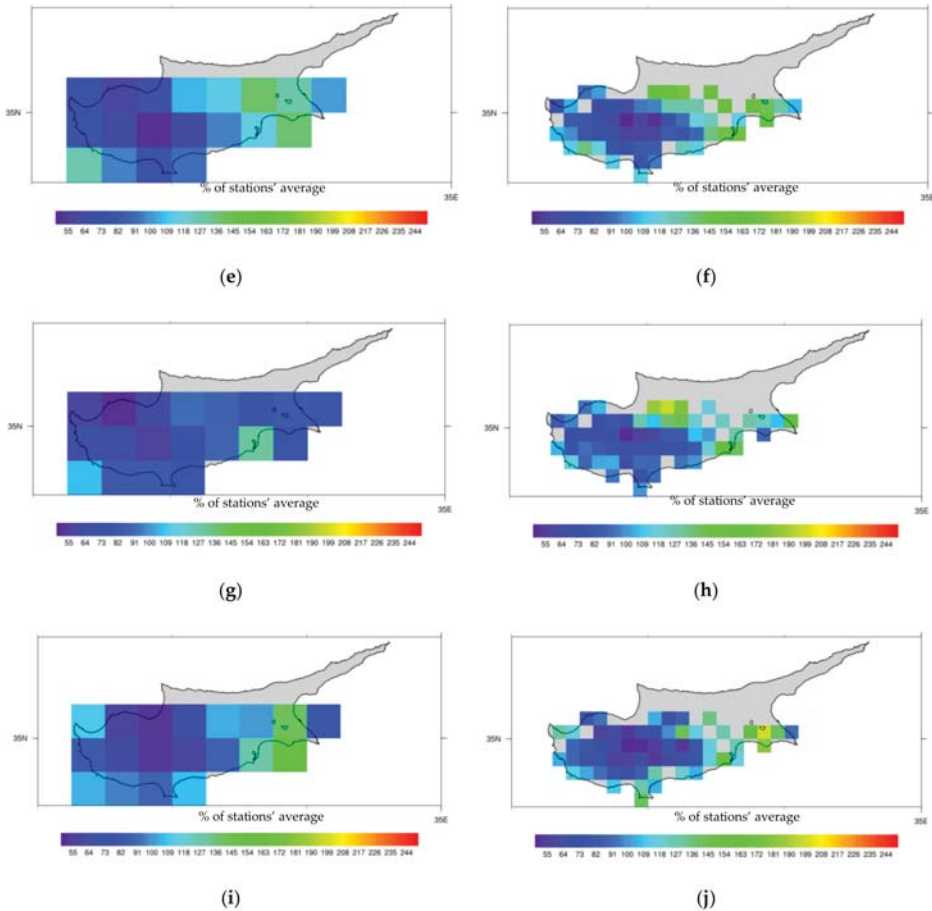
Figure 7. Underestimation/Overestimation of the annual satellite product estimates as compared with gauge station records at the corresponding grid cell: (**a**) TRMM—2014 (April–December); (**b**) GPM—2014 (April–December); (**c**) TRMM—2015; (**d**) GPM—2015; (**e**) TRMM—2016; (**f**) GPM—2016; (**g**) TRMM—2017; (**h**) GPM—2017; (**i**) TRMM—2018 (January–June); (**j**) GPM—2018 (January–June).

The performance of the IMERG product was slightly better than 3B43V7 with an R higher value (0.82 and 0.81, respectively), lower Bias values (-17.9 mm/month and -30.5 mm/month, respectively) and rBias values (3.1 and -3.0, respectively), while presenting worst performance in terms of RMSE (108.8 mm/month and 94.7 mm/month, respectively) and MAE (96.1 mm/month and 83.4 mm/month, respectively) values, as presented in Table 4 (minimum and maximum respective values are provided in brackets).

Table 4. Performance evaluation metrics of annual data from 3B43V7 and IMERG and (the range of values is given in parentheses).

	TRMM 3B43	GPM IMERG
R (correlation coefficient)	0.81 (0.46–0.99)	0.82 (0.37–0.99)
Mean Error (Bias)	−30.5 (−219.3–77.0)	−17.9 (−324.4–132.2)
Relative Bias	−3.0 (−38.6–35.0)	3.1 (−45.6–58.7)
RMSE	94.7 (37.5–234.8)	108.8 (32.3–335.0)
Mean Absolute Error	83.4 (30.2–219.3)	96.1 (25.2–324.4)

4.4. Influence of Elevation on Satellite Precipitation Products

To analyze further the influence of elevation and satellite precipitation products, we grouped all the rain gauge stations into three categories according to their elevation (0–300 m, 300–600 m, >600 m), and compared the evaluation metrics across the different elevation ranges. The annual rBias results for the both GPM and TRMM products are presented in Tables 5 and 6, respectively. Minimum and maximum respective values are provided in brackets along with the corresponding number of satellite products cell for each elevation category.

Results as presented in Tables 5 and 6, respectively, portray that, on an annual scale, both of the two SP products overestimate the precipitation below an altitude of 300 m, with IMERG presenting the largest overestimation (mean annual RB values: RBIMERG = 123.8%, RB3B43 = 107.7%). On the contrary, both SP products underestimate precipitation with increasing elevation, with 3B43v7 displaying a more apparent underestimation than IMERG, with RBIMERG = 93.0%, RB3B43 = 78.0% and RBIMERG = 70.3%, RB3B43 = 65.6% for elevation ranges between 300 and 600 m and >600 m, respectively. Regarding the Pearson correlation coefficient, the performance for 3B43 v7 was slightly better than IMERG at all of the categories, with *r* values increasing with elevation.

Table 5. Annual rBias and Pearson correlation coefficient performance for GPM data according to elevation (the range of values is given in parentheses).

	All Annual	2014	2015	2016	2017	2018
% of gauge	123.8 (90.9–158.7)	133.8 (93.6–234.0)	133.8 (86.5–185.2)	127.3 (79.7–175.1)	118.1 (76.4–196.8)	119.6 (77.7–200.2)
correlation	0.90 (0.78–0.95)	Elevation: 0–300 m (30 cells)				
% of gauge	93.0 (68.5–131.1)	96.3 (68.8–144.7)	107.1 (74.5–145.2)	89.8 (58.5–127.9)	93.7 (68.5–145.4)	81.9 (52.6–143.4)
correlation	0.91 (0.85–0.96)	Elevation: 300–600 m (19 cells)				
% of gauge	70.3 (54.4–81.3)	77.9 (61.8–109.4)	82.7 (64.9–101.6)	65.8 (50.3–92.9)	68.8 (50.3–85.8)	58.3 (44.0–72.8)
correlation	0.92 (0.88–0.94)	Elevation: >600 m (12 cells)				

Table 6. Annual rBias and Pearson correlation coefficient performance for TRMM data according to elevation (the range of values is given in parentheses).

	All Annual	2014	2015	2016	2017	2018
% of gauge	107.7 (85.9–135.0)	119.2 (91.9–166.8)	125.5 (86.7–182.6)	106.4 (77.4–139.3)	85.9 (61.6–131.8)	107.8 (82.5–160.7)
correlation	0.92 (0.84–0.96)	Elevation: 0–300 m (13 cells)				
% of gauge	78.0 (67.3–99.6)	83.1 (73.8–102.6)	99.3 (83.9–122.1)	74.8 (61.1–106.9)	68.1 (51.9–91.3)	60.6 (52.6–66.7)
correlation	0.92 (0.88–0.94)	Elevation: 300–600 m (4 cells)				
% of gauge	65.6 (61.4–69.8)	74.1 (72.1–76.1)	77.3 (75.9–78.7)	57.2 (51.2–63.1)	63.1 (57.2–69.0)	56.5 (52.6–60.5)
correlation	0.94 (0.93–0.94)	Elevation: >600 m (2 cells)				

More evaluation metrics (Bias, RMSE, MAE) were used to evaluate the performance of GPM and TRMM monthly data according to elevation (Table 7). Overall, the performance of both satellite products metrics (Bias, RMSE, MAE) were worst in higher altitude areas than in lower altitude areas. Regarding Bias, findings established that TRMM performed better in the elevation range 0–300 m, while GPM exhibited lower bias values in higher altitudes. TRMM exhibited lower RMSE and MAE values than those of GPM in the elevation ranges 0–300 m and >600 m and higher values in the elevation range 300–600 m.

Table 7. Metrics (Bias, RMSE, MAE) performance for GMP and TRMM monthly data according to elevation (the range of values is given in parentheses).

Elevation (m)	Bias (mm/month)		RMSE (mm/month)		MAE (mm/month)	
	GPM	*TRMM*	*GPM*	*TRMM*	*GPM*	*TRMM*
0–300	5.8 (−5.2–13.0)	1.5 (−5.8–7.7)	19.6 (14.7–28.6)	17.2 (11.6–25.1)	13.2 (10.2–17.2)	11.0 (9.0–14.3)
300–600	−3.9 (−17.0–8.1)	−10.2 (−15.8—0.1)	23.1 (15.4–42.1)	24.5 (13.4–30.0)	14.3 (10.3–21.8)	15.3 (8.7–18.0)
>600	−16.5 (−31.8—8.7)	−17.8 (−21.5—14.1)	35.4 (23.2–55.0)	33.0 (28.7–37.2)	21.5 (15.0–35.5)	21.0 (17.7–24.3)

The monthly and annual spatio-temporal variations of bias for both satellite precipitation products for the study period are presented in Figures 8 and 9, respectively, while the corresponding seasonal spatio-temporal variation is presented in Figure 10. It is clear that both SP products underestimate precipitation in higher elevation areas and overestimate in areas with lower elevation fluctuations. The underestimation is more evident in the winter.
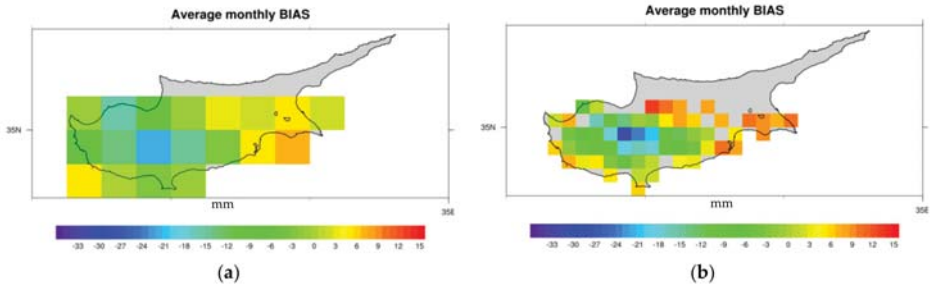
Figure 8. Monthly bias for both precipitation products for the study period: (**a**) TRMM and (**b**) GPM.
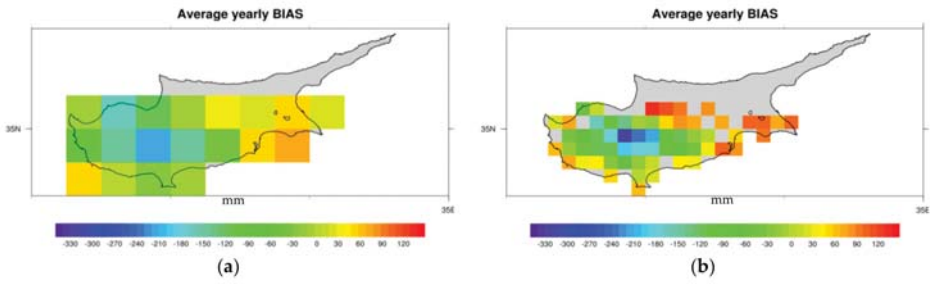
Figure 9. Annual bias for both precipitation products for the study period: (**a**) TRMM and (**b**) GPM.
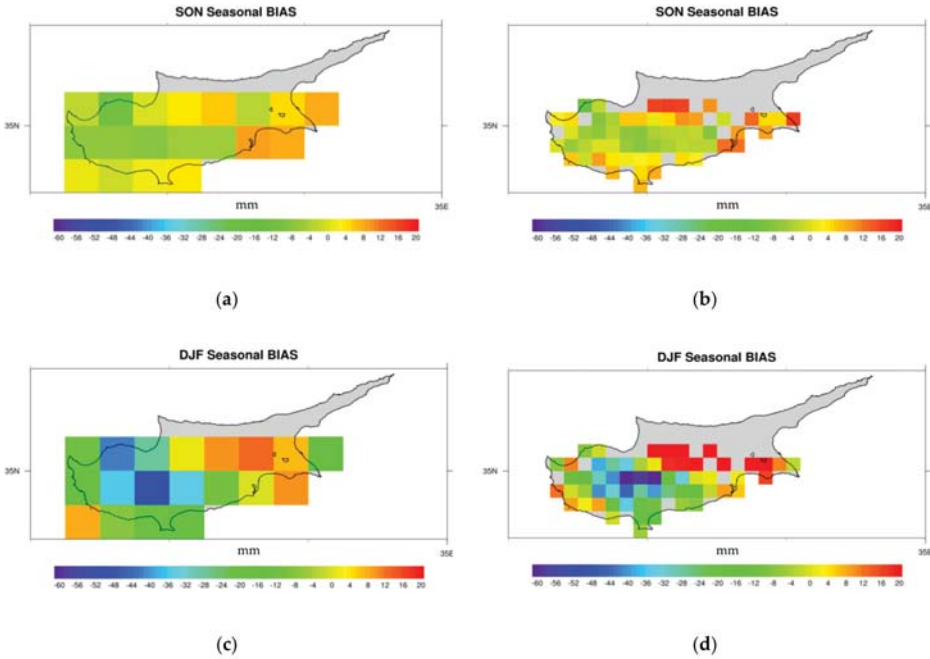
Figure 10. *Cont.*
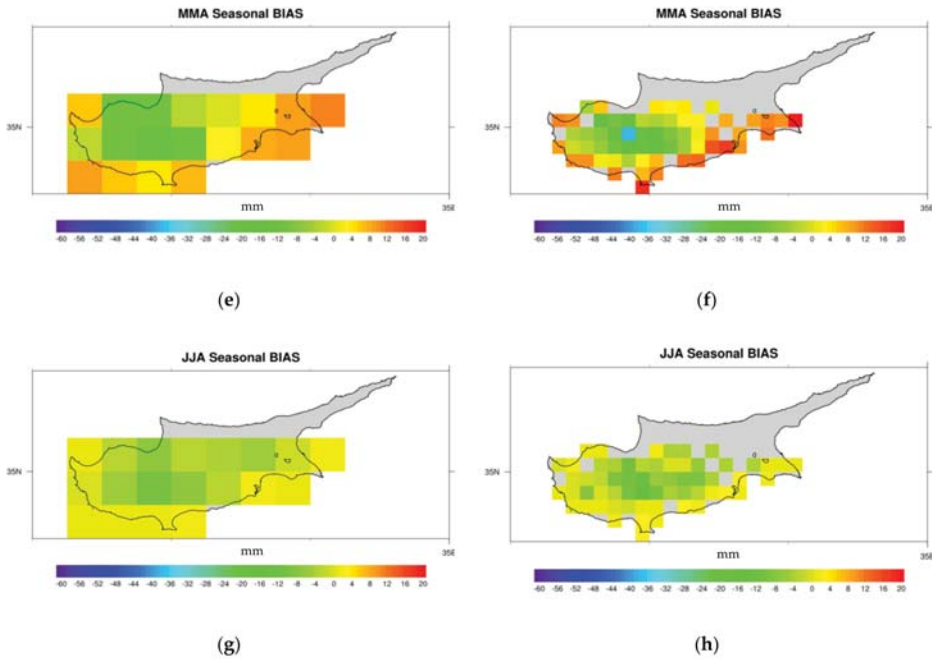
Figure 10. Seasonal bias for both precipitation products for the study period: (**a**) TRMM for September, October and November; (**b**) GPM for September, October and November; (**c**) TRMM for December, January and February; (**d**) GPM for December, January and February; (**e**) TRMM for March, April and May; (**f**) GPM for March, April and May; (**g**) TRMM for June, July and August; (**h**) GPM for June, July and August.

These findings are in agreement with previous studies reported in the literature [45–50]. The performance of both SP products could be due to the products themselves and to topography. Both satellite precipitation products combine data from both satellite sensors and ground gauges. Since data from only one gauge station are used in mountainous areas in Cyprus, while three are located in rather flat areas, the accuracy of satellite precipitation products may be affect. Chen and Li [51] and Tang et al. [52] also reported that the accuracy of satellite precipitation products in high mountainous areas in west China could be attributed to the sparse gauge network. Moreover, estimated differences could be also attributed to the differences of the rainfall process, which is rather complicated in mountainous areas than in low altitude areas due to the influence of topography.

5. Conclusions

An evaluation of the monthly and annual IMERG and TRMM 3B43 product estimates with corresponding rain-gauges data over Cyprus for the period April 2014 to June 2018 was performed. Based on the analysis presented, it is found that, overall, both monthly satellite product estimates and rain gauge data presented a very good agreement; however, both IMERG and TRMM estimates tend to underestimate precipitation, especially during the rainy season, although, IMERG and rain gauge records seem to exhibit similar temporal patterns. Considering the annual values, we notice that both SP estimates underestimate annual precipitation records, although IMERG estimates are much closer to gauge station records. In terms of statistical scores analysis, it was found that on a monthly and annual basis, a slightly better performance of IMERG for R, Bias and rBias values was noticed, while 3B43 product performed better in terms of RMSE and MAE values. Seasonal analysis showed that both products exhibited a better performance during the rainy (winter) period, followed by autumn and

spring seasons, while both products were able to detect the summer-time precipitation, although with high uncertainty in terms of relative bias values.

In summary, we conclude that although satellite products could be considered as quite accurate estimates of precipitation, indeed, their accuracy is not yet profound, and this issue is open to further elaboration. Nonetheless, IMERG estimates, due to their superiority in terms of spatial and temporal resolution, could serve as an alternative precipitation dataset, where in-situ precipitation records are limited.

The influence of elevation of both SP estimates was considered by grouping rain gauge stations in three categories, with respect to their elevation and it was found that both SP estimates underestimate precipitation with increasing elevation and overestimate it at lower elevations. Thus, it is suggested that one possible improvement would be the prospect of blending the SP data with more in situ data from rain gauges that are distributed evenly over the geographical area of Cyprus and especially in mountainous areas. Furthermore, it would be quite challenging to enhance the retrieval algorithms by implementing elevation correction or adjustment.

Although the results derived from this study are site specific for Cyprus, the methodology adopted could be "transferred" to other regions according to our understanding of how satellite-based precipitation estimates perform over different regions. For example, for study areas with characteristics similar to our area of study, in terms of geographic location, with no very complex topography, the methodology could be applied directly. For other areas, with complex topography or with various climatic zones the methodology should further consider these parameters.

The authors aim to expand their research in this field, by considering an evaluation over Cyprus of the newly released version of IMERG data that expands the SP products into a uniformly processed data set embracing the TRMM-era. The establishment of a uniform TRMM and GPM SP record will broaden the scientific challenges for further research in several meteorological and hydrological applications.

References

1. Yilmaz, K.K.; Hogue, T.S.; Hsu, K.-L.; Sorooshian, S.; Gupta, H.V.; Wagener, T. Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeorol.* **2005**, *6*, 497–517. [CrossRef]
2. Shrestha, M.S.; Artan, G.A.; Bajracharya, S.R.; Sharma, R.R. Using satellite-based rainfall estimation for stream-flow modeling: Bagmati basin. *J. Flood Risk Manag.* **2008**, *1*, 88–89. [CrossRef]
3. Tobin, K.J.; Bennett, M.E. Using SWAT to model streamflow in two river basins with ground and satellite precipitation data. *J. Am. Water Resour. Assoc.* **2009**, *45*, 253–271. [CrossRef]
4. Bui, H.T.; Ishidaira, H.; Shaowei, N. Evaluation of the use of global satellite–gauge and satellite-only precipitation products in stream flow simulations. *Appl. Water. Sci.* **2019**, *9*, 53. [CrossRef]
5. Akinyemi, D.F.; Ayanlade, O.S.; Nwaezeigwe, J.O. A Comparison of the accuracy of multi-satellite precipitation estimation and ground meteorological records over southwestern Nigeria. *Remote Sens. Earth Syst. Sci.* **2019**. [CrossRef]

6. Levizzani, V.; Cattani, E. Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. *Remote Sens.* **2019**, *11*, 2301. [CrossRef]

7. Prigent, C. Precipitation retrieval from space: An overview. *C. R. Geosci.* **2010**, *342*, 380–389. [CrossRef]

8. Kummerow, C.D.; Randel, D.L.; Kulie, M.; Wang, W.; Ferraro, R.; Joseph Munchak, S.; Petkovic, V. The evolution of the Goddard profiling algorithm to a fully parametric scheme. *J. Atmos. Ocean. Technol.* **2015**, *32*, 2265–2280. [CrossRef]

9. Tapiador, F.J.; Navarro, A.; Levizzani, V.; García-Ortega, E.; Huffman, G.J.; Kidd, C.; Kucera, P.; Kummerow, C.D.; Masunaga, H.; Petersen, W.A.; et al. Global precipitation measurements for validating climate models. *Atmos. Res.* **2017**, *94*, 512–533. [CrossRef]

10. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K.-L. A review of global precipitation data sets: Data sources, estimation, and inter-comparisons. *Rev. Geophys.* **2018**, *56*, 79–107. [CrossRef]

11. Skofronick-Jackson, G.; Kulie, M.; Milani, L.; Munchak, S.J.; Wood, N.B.; Levizzani, V. Satellite Estimation of falling snow: A Global Precipitation Measurement (GPM) core observatory perspective. *J. Appl. Meteorol. Climatol.* **2019**, *58*, 1429–1448. [CrossRef]

12. Dias, J.; Gehne, M.; Kiladis, G.N.; Sakaeda, N.; Bechtold, P.; Haiden, T. Experimental assimilation of the GPM core observatory DPR reflectivity profiles for typhoon Halong (2014). *Mon. Wea. Rev.* **2016**, *144*, 2307–2326. [CrossRef]

13. Panegrossi, G.; Casella, D.; Dietrich, S.; Marra, A.C.; Sano, P.; Mugnai, A.; Baldini, L.; Roberto, N.; Adirosi, E.; Cremonini, R.; et al. Use of the GPM constellation for monitoring heavy precipitation events over the Mediterranean region. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2733–2753. [CrossRef]

14. Skofronick-Jackson, G.; Petersen, W.A.; Beng, W.; Kidd, C.; Stocker, E.F.; Kirschbaum, D.B.; Kakar, R.; Braun, S.A.; Huffman, G.J.; Iguchi, T.; et al. The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1679–1695. [CrossRef] [PubMed]

15. Marra, A.C.; Porcù, F.; Baldini, L.; Petracca, M.; Casella, D.; Dietrich, S.; Mugnai, A.; Sanò, P.; Vulpiani, G.; Panegrossi, G. Observational analysis of an exceptionally intense hailstorm over the Mediterranean area: Role of the GPM Core Observatory. *Atmos. Res.* **2017**, *192*, 72–90. [CrossRef]

16. Marra, A.C.; Federico, S.; Montopoli, M.; Avolio, E.; Baldini, L.; Casella, D.; D'Adderio, L.P.; Dietrich, S.; Sanò, P.; Torcasio, C.R.; et al. The Precipitation structure of the Mediterranean tropical-like cyclone Numa: Analysis of GPM Observations and Numerical Weather Prediction model simulations. *Remote Sens.* **2019**, *11*, 1690. [CrossRef]

17. Panegrossi, G.; Marra, A.C.; Sanò, P.; Baldini, L.; Casella, D.; Porcù, F. Heavy precipitation systems in the Mediterranean area: The role of GPM. In *Satellite Precipitation Measurement-Volume 1*; Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds.; Springer: Dordrecht, The Netherlands, 2020; pp. 819–841. ISBN 978-3-030-24567-2. [CrossRef]

18. Liu, Z.; Ostrenga, D.; Vollmer, B.; Deshong, B.; Macritchie, K.; Greene, M.; Kempler, S. Global precipitation measurement mission products and services at the NASA GES DISC. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 437–444. [CrossRef]

19. Wang, Z.; Zhong, R.; Lai, C.; Chen, J. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.* **2017**, *196*, 151–163. [CrossRef]

20. Wu, L.; Xu, Y.; Wang, S. Comparison of TMPA-3B42RT legacy product and the equivalent IMERG products over mainland China. *Remote Sens.* **2018**, *10*, 1778. [CrossRef]

21. Anjum, M.N.; Ahmad, I.; Ding, Y.; Shangguan, D.; Zaman, M.; Ijaz, M.W.; Sarwar, K.; Han, H.; Yang, M. Assessment of IMERG-V06 precipitation product over different hydro-climatic regimes in the Tianshan. *Remote Sens.* **2019**, *11*, 2314. [CrossRef]

22. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]

23. Wang, X.; Ding, Y.; Zhao, C.; Wang, J. Similarities and improvements of GPM IMERG upon TRMM 3B42 precipitation product under complex topographic and climatic conditions over Hexi region, northeastern Tibetan Plateau. *Atmos. Res.* **2019**, *218*, 347–363. [CrossRef]

24. Chen, C.; Chen, Q.; Duan, Z.; Zhang, J.; Mo, K.; Li, Z.; Tang, G. Multiscale comparative evaluation of the GPM IMERG v5 and TRMM 3B42 v7 precipitation products from 2015 to 2017 over a climate transition area of China. *Remote Sens.* **2018**, *10*, 944. [CrossRef]

25. Wei, G.; Lü, H.; Crow, W.T.; Zhu, Y.; Wang, J.; Su, J. Comprehensive evaluation of GPM-IMERG, CMORPH, and TMPA precipitation products with gauged rainfall over mainland China. *Adv. Meteorol.* **2018**, 3024190. [CrossRef]

26. Sunilkumar, K.; Yatagai, A.; Masuda, M. Preliminary evaluation of GPM-IMERG rainfall estimates over three distinct climate zones with APHRODITE. *Earth Space Sci.* **2019**, *6*, 1321–1335. [CrossRef]

27. Retalis, A.; Katsanos, D.; Michaelides, S. Precipitation climatology over the Mediterranean Basin–Validation over Cyprus. *Atmos. Res.* **2016**, *169*, 449–458. [CrossRef]

28. Katsanos, D.; Retalis, A.; Michaelides, S. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. *Atmos. Res.* **2016**, *169*, 459–464. [CrossRef]

29. Katsanos, D.; Retalis, A.; Tymvios, F.; Michaelides, S. Analysis of precipitation extremes based on satellite (CHIRPS) and in-situ data set over Cyprus. *Nat. Hazards* **2016**, *83*, S53–S63. [CrossRef]

30. Retalis, A.; Katsanos, D.; Tymvios, F.; Michaelides, S. Validation of the first years of GPM operation over Cyprus. *Remote Sens.* **2018**, *10*, 1520. [CrossRef]

31. Nicolaides, K.A.; Michalelides, S.C.; Karacostas, T. Synoptic and dynamic characteristics of selected deep depressions over Cyprus. *Adv. Geosci.* **2006**, *7*, 175–180. [CrossRef]

32. Nicolaides, K.; Savvidou, K.; Orphanou, A.; Michaelides, S.C.; Constantinides, P.; Papachristodoulou, C.; Savvides, M. An investigation of a baroclinic depression that affected the area of Cyprus. *Adv. Geosci.* **2008**, *16*, 117–124. [CrossRef]

33. Michaelides, S.; Tymvios, F.; Michaelidou, T. Spatial and temporal characteristics of the annual rainfall frequency distribution in Cyprus. *Atmos. Res.* **2009**, *94*, 606–615. [CrossRef]

34. Cyprus Department of Meteorology. The Climate of Cyprus. Available online: http://www.moa.gov.cy/moa/ms/ms.nsf/DMLcyclimate_en/DMLcyclimate_en? (accessed on 4 April 2020).

35. Hou, A.Y.; Skofronick-Jackson, G.; Kummerow, C.D.; Shepherd, J.M. Global Precipitation Measurement. In *Precipitation: Advances in Measurement, Estimation and Prediction*; Michaelides, S., Ed.; Springer: Berlin, Germany, 2008; pp. 131–169. ISBN 978-3-540-77654-3.

36. Gebregiorgis, A.S.; Kirstetter, P.-E.; Hong, Y.E.; Gourley, J.J.; Huffman, G.J.; Petersen, W.A.; Xue, X.; Shwaller, M. To what extent is the day 1 GPM IMERG satellite precipitation estimate improved as compared to TRMM TMPA-RT? *J. Geophys. Res. Atmos.* **2018**, *123*, 1694–1707. [CrossRef]

37. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Oki, R.; Nakamura, K.; Ihuchi, T. The Global Precipitation Measurement mission. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 701–722. [CrossRef]

38. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Integrated Multi-SatellitE Retrievals for GPM (IMERG) Technical Documentation*; NASA/GSFC: Greenbelt, MD, USA, 2015.

39. Elsaesser, G.S.; Kummerow, C.D. The Sensitivity of rainfall estimation to error assumptions in a Bayesian passive microwave retrieval algorithm. *J. Appl. Meteor. Climatol.* **2015**, *54*, 408–422. [CrossRef]

40. Kummerow, C.; Barnes, W.; Kozu, T.; Shiue, J.; Simpson, J. The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Ocean. Technol.* **1998**, *15*, 809–817. [CrossRef]

41. Kummerow, C.; Simpson, J.; Thiele, O.; Barnes, W.; Chang, A.T.C.; Stocker, E.; Adler, R.F.; Hou, A.; Kakar, R.; Wentz, F.; et al. The Status of the Tropical Rainfall Measuring Mission (TRMM) after 2 years in orbit. *J. Appl. Meteorol.* **2000**, *39*, 1965–1982. [CrossRef]

42. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.J.; Bowman, K.P.; Hong, Y.; Stocker, E.F.; Wolff, D.B. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]

43. Huffman, G.J. The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG) (TMPA to IMERG). Available online: https://gpm.nasa.gov/resources/documents/transition-multi-satellite-products-trmm-gpm-tmpa-imerg (accessed on 2 April 2020).

44. Munchak, S.J.; Skofronick-Jackson, G. Evaluation of precipitation detection over various surfaces from passive microwave imagers and sounders. *Atmos. Res.* **2013**, *131*, 81–94. [CrossRef]

45. Jin, X.L.; Shao, H.; Zhang, C.; Yan, Y. The applicability evaluation of three satellite products in Tianshan Mountains. *J. Nat. Resour.* **2016**, *31*, 2074–2085. [CrossRef]

46. Ma, Y.; Tang, G.; Long, D.; Yong, B.; Zhong, L.; Wan, W.; Hong, Y. Similarity and error intercomparison of the GPM and its predecessor-TRMM Multisatellite Precipitation Analysis using the best available hourly gauge network over the Tibetan Plateau. *Remote Sens.* **2016**, *8*, 569. [CrossRef]

47. Xu, R.; Tian, F.; Yang, L.; Hu, H.; Lu, H.; Hou, A. Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.* **2017**, *122*, 910–924. [CrossRef]

48. Fang, J.; Yang, W.; Luan, Y.; Du, J.; Lin, A.; Zhao, L. Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. *Atmos. Res.* **2019**, *223*, 24–38. [CrossRef]

49. Wang, S.; Liu, J.; Wang, J.; Qiao, X.; Zhang, J. Evaluation of GPM IMERG V05B and TRMM 3B42V7 precipitation products over High Mountainous tributaries in Lhasa with dense rain gauges. *Remote Sens.* **2019**, *11*, 2080. [CrossRef]

50. Wu, Y.; Zhang, Z.; Huang, Y.; Jin, Q.; Chen, X.; Chang, J. Evaluation of the GPM IMERG v5 and TRMM 3B42 v7 precipitation products in the Yangtze river basin, China. *Water* **2019**, *11*, 1459. [CrossRef]

51. Chen, F.R.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]

52. Tang, G.; Zeng, Z.; Long, D.; Guo, X.; Yong, B.; Zhang, W.; Hong, Y. Statistical and hydrological comparisons between TRMM and GPM level-3 products over a midlatitude basin: Is day-1 IMERG a good successor for TMPA 3B42V7? *J. Hydrometeorol.* **2016**, *17*, 121–137. [CrossRef]

GPM-Based Multitemporal Weighted Precipitation Analysis Using GPM_IMERGDF Product and ASTER DEM in EDBF Algorithm

Sana Ullah [1,2], Zhengkang Zuo [1], Feizhou Zhang [1], Jianghua Zheng [3], Shifeng Huang [4], Yi Lin [1], Imran Iqbal [5], Yiyuan Sun [1,6], Ming Yang [1] and Lei Yan [1,2,*]

[1] Beijing Key Laboratory of Space Information Integration and 3s Application, School of Earth and Space Science, Peking University, Beijing 100871, China; sana_ullah@pku.edu.cn (S.U.); 1801110646@pku.edu.cn (Z.Z.); zhangfz@pku.edu.cn (F.Z.); yi.lin@pku.edu.cn (Y.L.); sunyiyuan@pku.edu.cn (Y.S.); yangming95@buaa.edu.cn (M.Y.)

[2] Guangxi Key Laboratory of Remote Measuring System, Guilin University of Aerospace Technology, Guilin 541004, China

[3] College of Resources and Environment Science, Xinjiang University, Urumqi 830046, China; zheng.jianghua@xju.edu.cn

[4] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China; huangsf@iwhr.com

[5] Department of Information and Computational Sciences, School of Mathematical Sciences and LMAM, Peking University, Beijing 100871, China; imraniqbalrajput@pku.edu.cn

[6] Key Laboratory of Mountain Resources and Environmental Remote Sensing, School of Geography and Environmental Science, Guizhou Normal University, Guiyang 550001, China

* Correspondence: lyan@pku.edu.cn

Abstract: To obtain the high-resolution multitemporal precipitation using spatial downscaling technique on a precipitation dataset may provide a better representation of the spatial variability of precipitation to be used for different purposes. In this research, a new downscaling methodology such as the global precipitation mission (GPM)-based multitemporal weighted precipitation analysis (GMWPA) at 0.05° resolution is developed and applied in the humid region of Mainland China by employing the GPM dataset at 0.1° and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) 30 m DEM-based geospatial predictors, i.e., elevation, longitude, and latitude in empirical distribution-based framework (EDBF) algorithm. The proposed methodology is a two-stepped process in which a scale-dependent regression analysis between each individual precipitation variable and the EDBF-based weighted precipitation with geospatial predictor(s), and to downscale the predicted multitemporal weighted precipitation at a refined scale is developed for the downscaling of GMWPA. While comparing results, it shows that the weighted precipitation outperformed all precipitation variables in terms of the coefficient of determination (R^2) value, whereas they outperformed the annual precipitation variables and underperformed as compared to the seasonal and the monthly variables in terms of the calculated root mean square error (RMSE) value. Based on the achieved results, the weighted precipitation at the low-resolution (e.g., at 0.75° resolution) along-with the original resolution (e.g., at 0.1° resolution) is employed in the downscaling process to predict the average multitemporal precipitation, the annual total precipitation for the year 2001 and 2004, and the average annual precipitation (2001–2015) at 0.05° resolution, respectively. The downscaling approach resulting through proposed methodology captured the spatial patterns with greater accuracy at higher spatial resolution. This work showed that it is feasible to increase the spatial resolution of a precipitation variable(s) with greater accuracy on an annual basis or as an average from the multitemporal precipitation dataset using a geospatial predictor as the proxy of precipitation through the weighted precipitation in EDBF environment.

1. Introduction

Precipitation is the major component of the global water cycle. It is a key parameter of the ecological, hydrological, meteorological and agriculture systems [1,2]. It plays an important role in the energy exchange and material circulation of the Earth surface system [3]. It is of significant importance to understand the characteristics of precipitation, because it shows great variability both in space and time as compared to other climatic variables. Therefore, its spatial and temporal variability greatly influence vegetation distribution, soil moisture and surface runoff [4,5]. In addition, a high-quality precipitation dataset is very important in the development of different ecological and hydrological models at corresponding scales. On top of that, due to certain limiting factors, it is difficult to develop such high-quality dataset(s) from point measurements based on the traditional precipitation, which are as follows: first, the data derived from point measurements heavily depends on field observations [5,6]. Second, field observation stations are not uniformly distributed in space and limited mostly to low and medium altitude areas, with the exception of a few precipitation stations at high altitudes. Moreover, their operational capability is relative for a shorter period. Even if longer precipitation records exist from ground-based stations, they are not sufficient to provide coverage for the global/regional applications, due to deficiencies in reliability of the spatial distribution of precipitation [7], especially over ocean, desert and mountainous areas. Third, a true spatial coverage of precipitation based on the traditional rain gauge observations cannot be obtained [8], because many river basins around the world are still poorly gauged [9], or ungauged [10]. Fourth, it is difficult to effectively reflect the spatial variability of precipitation based on the observation from a finite number of rainfall stations, especially in areas where rainfall stations are sparsely distributed [11–13]. Fifth, rain gauge observations can only reflect the point rainfall within a radius around the location of instruments, and the effectiveness of such data is often under question, and adequate validation is further needed [14,15].

Recently, the development in remote sensing and geographic information technology has given a new dimension to present precipitation observations [16–18], almost at the global scale over a long period, which also reflects the spatial patterns and temporal variability of precipitation [19]. In this regard, various research institutions and government organizations have developed a series of gridded global precipitation datasets, including Earth observations, in situ datasets and models at both regional and global scales, i.e., the Global Precipitation Climatology Project (GPCP) [2,20–22], the Global Satellite Mapping of Precipitation (GSMaP) project [23], the Multi-Source Weighted-Ensemble Precipitation (MSWEP) [24], the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) [25], the Precipitation Estimation from Remotely-Sensed Information using Artificial Neural Networks-Climate Record (PERSIANN-CDR) [26], the Tropical Rainfall Measuring Mission (TRMM) [27–29], the TRMM Multi-satellite Precipitation Analysis (TMPA) [30], and the Global Precipitation Mission (GPM) [31–33].

Spatial downscaling is a recently developed approach to obtain the high spatial resolution of a variable based on conjugation between the variable at a coarse scale and geospatial predictor(s) at the low resolution [34,35]. In this regard, using spatial downscaling techniques on a precipitation dataset may provide a better representation of the spatial variability of precipitation to be used for different purposes. Several authors have used downscaling methodologies to increase the spatial resolution of satellite-based precipitation, often in combination with Earth observations data available on hydro-meteorological variables related to precipitation, including normalized difference vegetation index (NDVI) [30,35–42], digital elevation model [30,38,43,44], land surface temperature [30], soil moisture [37], in situ rain gauged precipitation [37,38], slope [38], aspect [38], and wind [31]. Moreover, few authors have used different satellite-based precipitation datasets for

TRMM products [30,42]. Additionally, some studies used regression analysis with model parameters spatially constant (multiple linear, polynomial, exponential, regression kriging, etc.), assuming a spatial stationarity of the relationship between precipitation and the predicting variables [34,35,38,41,44–48]. On top of that, some studies limited their analysis only to satellite-based precipitation datasets and did not take full advantage of all available data sources, combining remotely sensed and in situ observations [42,49,50].

In this research work, a new downscaling methodology (Figure 1), based on the earlier work of [3,38,39], such as the GMWPA is developed using DEM (Figure 2a) to delineate into three geospatial predictors, i.e., elevation, longitude, and latitude [44], in EDBF algorithm. Two different satellite-based precipitation datasets, such as the GPM-based multitemporal precipitation data (Figure 2b–i) for the prediction of high-resolution downscaled weighted precipitation from 0.1° to 0.05° resolution, and the GPM (Figure 2j,k) and the TRMM (Figure 2l–o) datasets for the verification of proposed methodology is used over the humid (the Southern) region of Mainland China. During the execution, certain objectives are set to achieve the required results, which are as follows [51]: to evaluate the multitemporal precipitation (2001–2015) dataset through regression analysis, i.e., polynomial regression at different upscaled resolutions, e.g., 0.25°, 0.5°, 0.75°, 1.0°, 1.25°, 1.50°; (2) based on the regression output, EDBF algorithm is run to evaluate the multitemporal precipitation at each upscaled resolution by assigning weight to each temporal component; (3) to verify the output of EDBF algorithm through the TRMM and the GPM datasets; and (4) to generate the high-resolution downscaled weighted precipitation at 0.05° resolution based on the best performing upscaled resolution. This research can have practical implications, particularly for climate change, drought assessment, and water resources planning, which require long-term precipitation estimates at finer resolution.
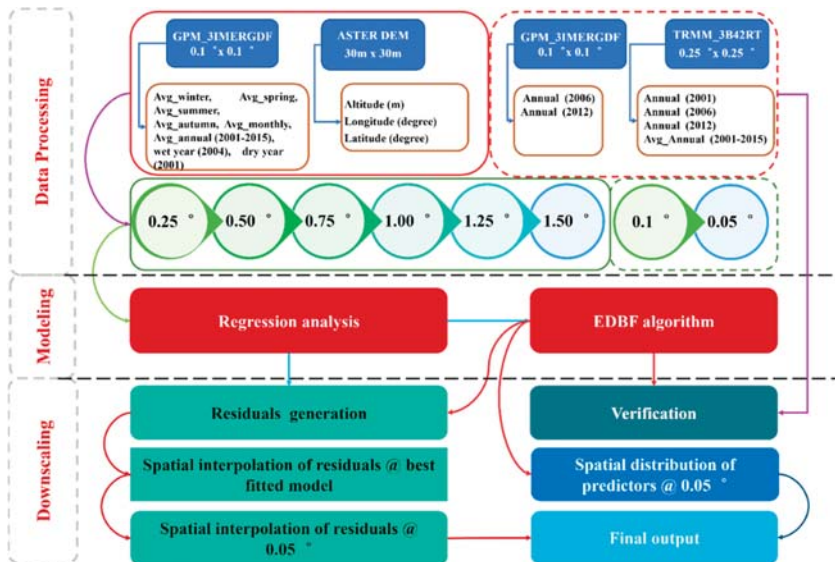
Figure 1. Proposed methodology to predict the high-resolution downscaled weighted precipitation.
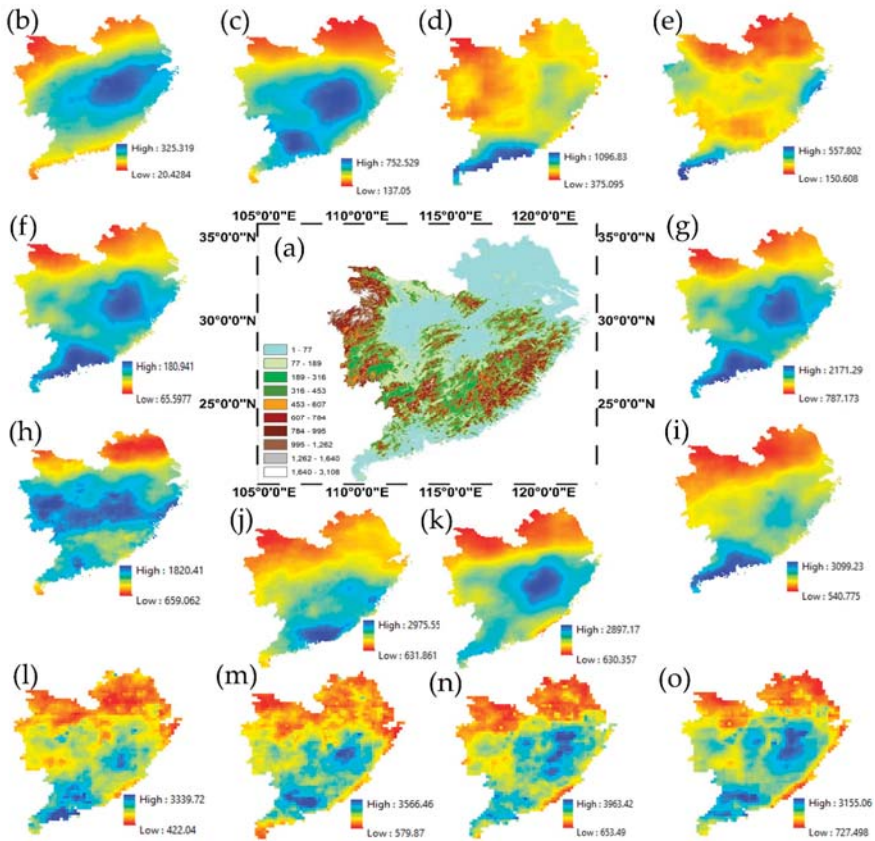
Figure 2. The dataset required for downscaling of the multitemporal precipitation: (**a**) the DEM of study area; (**b**) the GPM-based average winter precipitation; (**c**) the GPM-based average spring precipitation; (**d**) the GPM-based average summer precipitation; (**e**) the GPM-based average autumn precipitation; (**f**) the GPM-based average monthly precipitation; (**g**) the GPM-based average annual (2001–2015) precipitation; (**h**) the GPM-based wet year (2004) precipitation; (**i**) the GPM-based dry year (2001) precipitation; (**j**) the GPM-based annual (2006) precipitation; (**k**) the GPM-based annual (2012) precipitation; (**l**) the TRMM-based annual (2001) precipitation; (**m**) the TRMM-based annual (2006) precipitation; (**n**) the TRMM-based annual (2012) precipitation; (**o**) the TRMM-based average annual (2001–2015) precipitation, respectively.

2. Materials and Methods

2.1. Study Area

This research is conducted over the Southern part of China, which in Chinese is simply referred to as "the South". The study area comprises of eight provinces (Anhui, Hunan, Hubei, Jiangxi, Jiangsu, Zhejiang, Fujian, and Guangdong) and one municipality (Shanghai) (Figure 2a). It is approximately the mega-region within China separated by the Qingling-Huaihe Line, which is a reference line used by geographers to distinguish between the Northern and the Southern China, corresponding roughly to 33rd parallel [52]. From Qingling-Huaihe, "Qingling" refers to the Qingling Mountain, and "Huaihe" refers to the Huai River, running from the Qingling Mountain in the West to the Huai River in the East. It divides Eastern China into the North and the South regions, which differ from each other in climate, demography and terrain. All major rivers of China flow across this region, e.g., the Yangtze River,

the Huai River, the Han River, the Qu River, the Qiangtang River, the Ou River, the Gan River, the Min River, the Xiang River, the Zi River, the Yuan River and the Lishui River. In addition, some major lakes such as the Dongting Lake, the Tai Lake and the Chaozhou Lake are also located in this region. Moreover, regions laying in the South of the line tend to be tropical and subtropical. Some major mountainous series such as the Huangshan, the Dabie, the She, the Mufu, the Jiuling, the Luoxiao, the Huaiyu, the Wuyi, the Jiulian, the Dayu, the Nong are also located in this region. The Southern part is hotter and wetter than the Northern part. Normally, the weather conditions are with short, cool, damp winters, and very hot, humid summers. The average temperature in winter and summer remains between 3 °C to 9 °C and 27 °C to 30 °C, respectively. The average annual precipitation is between 1200 to 1900 mm, much of it falling in the form of heavy rains occurring in late spring and summer. In addition, half of the most developed tier 1 cities of China are located in the South.

2.2. Datasets

2.2.1. Global Precipitation Mission

During the execution of present study, the Integrated Multi-Satellite Retrievals for GPM (IMERG), an algorithm that provides the multi-satellite precipitation for GPM users, is used to retrieve the required GPM data [33]. Specifically, the daily rainfall (mm day^{-1}) estimate from the GPM Level 3 IMERG *Final* Daily 10 × 10 km Version 06 (GPM_3IMERGDF_V06) is used as primary data, moreover, which is derived from half-hourly GPM_3IMERGHH available at https://giovanni.gsfc.nasa.gov/giovanni/ #dataKeyword=IMERGDF (Accessed on 20 June 2020). Besides, the derived result represents the final estimate of the daily accumulated precipitation combined with microwave-infrared. The dataset is produced at the National Aeronautics and Space Administration (NASA) [53], Goddard Earth Sciences [10], Data and Information Services Center (DISC) by simply summing the valid precipitation retrievals for the day in GPM_3IMERGHH and giving the result in (mm). The reason for selecting GPM _IMERG is attributed to following reasons: (i) satellite rainfall estimates with finest gridded data @ 0.1° × 0.1° spatial resolution, high temporal data supply, ranging from half-hourly to daily and monthly, the GPM system provides almost near real-time data with spatial coverage (−180.0, −90.0, 180.0, 90.0) and temporal coverage (2000–06–01 to 2020–03–01). Further details about the GPM _IMERGDF can be found at https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06/summary?keywords=GPM. (Accessed on 20 June 2020). The daily GPM_3IMERGDF product from 2001 to 2015 is retrieved for the study area into following multitemporal aggregates:

Average Seasonal Precipitation

The daily GPM_3IMERGDF product from 2001 to 2015 is aggregated into the average seasonal precipitation @ 0.1° spatial resolution as shown in Figure 2b–e. The equation deriving the average seasonal precipitation is as follows in Equation (1):

$$P_{GPM\_S} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} P_{GPM_{ij}}}{N} \tag{1}$$

where P_{GPM_S} is the average seasonal precipitation which individually corresponds to DJF (December, January and February), MAM (March, April and May), JJA (June, July and August), and SON (September, October and November), respectively. $P_{GPM_{ij}}$ is the daily GPM_3IMERGDF precipitation for *i*-th day, i.e., DJF (*n* = 90), MAM (*n* = 92), JJA (*n* = 92), SON (*n* = 91) and *j*-th year (*m* = 15), and *N* is the total number of observations. Hence, hereafter, the average winter precipitation for DJF, the average spring precipitation for MAM, the average summer precipitation for JJA, and the average autumn precipitation for SON will be used.

Average Monthly, and Average Annual Precipitation

The daily GPM_3IMERGDF product from 2001 to 2015 is aggregated into the average monthly and the average annual precipitation @ 0.1° spatial resolution as shown in Figure 2f,g, respectively. The equation deriving the average monthly, and the average annual precipitation is as follows in Equation (2):

$$P_{GPM\_Avg} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} P_{GPM_{ij}}}{N} \tag{2}$$

where P_{GPM_Avg} is the average monthly and the average annual precipitation for the study area, $P_{GPM_{ij}}$ is the daily GPM_3IMERGDF precipitation for *i*-th day ($n = 365$) and *j*-th year ($m = 15$), and N is the total number of observations (e.g., for the average monthly precipitation $N = 180$, and for the average annual precipitation $N = 15$).

Annual Total Precipitation

The daily GPM_3IMERGDF product (365 days) is aggregated into the annual total precipitation for the wet year (2004) and the dry year (2001), and also for the year 2006 and 2012 (i.e., for the verification) @ 0.1° spatial resolution as shown in Figure 2h–k, respectively. The equation deriving the annual total precipitation for the mentioned temporal periods is as follows in Equation (3):

$$P_{GPM_i} = \sum_{j=1}^{n} P_{GPM_{daily-j}} \tag{3}$$

where P_{GPM_i} is *i*-th annual precipitation (i.e., 2001, 2004, 2006 and 2012,) for the study area, $P_{GPM_{daily-j}}$ is the daily GPM_3IMERGDF precipitation for *j*-th day ($n = 365$).

2.2.2. ASTER Global Digital Elevation Model (GDEM)

The ASTER GDEM is released by the Ministry of Economy, Trade and Industry [53] of Japan, and the NASA [38,54]. The ASTER GDEM covers the land surface between 83°N and 83°S of the Earth, which includes the entire area of Mainland China. The latest GDEM version 3 was released on August 5, 2019, which added additional stereo-pairs with improved coverage and reducing the occurrence of artifacts. The refined production algorithm provides an improved spatial resolution, increased the horizontal and vertical accuracy [53]. It provides the spatial resolution of one arc-second (approximately 30 m), and is used in this study as an influencing predictor on precipitation (Figure 2a). In addition to the elevation data, two other terrain attributes, i.e., longitude and latitude, are also derived from the ASTER GDEM. The ASTER GDEM data is available at https://search.earthdata.nasa.gov/search/granules (Accessed on 20 June 2020).

2.2.3. Tropical Rainfall Measuring Mission

The Tropical Rainfall Monitoring Mission (TRMM) is a joint project between the NASA and the Japan Aerospace Exploration Agency (JAXA). The TRMM was launched on 27th November, 1997 [20,29,55]. It provides measurement for the intensity and areal coverage (60°S to 60°N) of tropical and subtropical precipitation, which covers about two third of the world's rainfall [38]. There is a range of orbital and gridded TRMM products available, i.e., 3B42RT and 3B43RT datasets [39]. Specifically, the daily rainfall (mm day^{-1}) estimate from the TRMM (TMPA-RT) Near Real-Time Precipitation L3 1day 0.25° × 0.25° Version 7 (TRMM_3B42RT_Daily) is used as the primary dataset during the verification process in the present study, moreover, which is derived from the original three-hour averaged precipitation values available at https://giovanni.gsfc.nasa.gov/giovanni/#service=TmAvMp&starttime=&endtime=&dataKeyword=TRMM (Accessed on 20 June 2020). Further details can be found at https://disc.gsfc.nasa.gov/datasets/TRMM_3B42RT_Daily_7/summary (Accessed on 20 June 2020). The daily TRMM_3B42RT product (365 days) is aggregated into the annual total precipitation for the

year 2001, 2006 and 2012 (Figure 2l–n) and the average annual (2001–2015) precipitation (Figure 2o) @ 0.25° spatial resolution, which is used during the verification process. The equation deriving the annual total and the average annual precipitation is given in Equations (2) and (3), respectively.

2.3. Methodology

In this research, a new downscaling methodology (Figure 1) based on the weighted precipitation is presented, at a regional scale, to downscale the multitemporal GPM data in the humid region of Mainland China. To execute the proposed algorithm, a two-stepped methodology is developed to successfully predict and downscale the investigated precipitation variables at a finer scale: first, to evaluate the relationship between each individual precipitation variable and the EDBF-based weighted precipitation with geospatial predictor(s) through regression models; and second, to downscale the predicted multitemporal weighted precipitation at a refined scale.

2.3.1. Pre-Processing of DEM and GPM Datasets

Pre-processing is carried out by extracting geospatial variables, i.e., elevation, longitude, and latitude at 30 m resolution and the GPM-derived precipitation variables, i.e., the average winter, the average spring, the average summer, the average autumn, the average monthly, the average annual (2001–2015), the wet year (2004), and the dry year (2001) precipitation at 0.1° resolution into six different resolution scales (i.e., 0.25°, 0.5°, 0.75°,1°, 1.25°and 1.50°, respectively) by applying the pixel averaging, e.g., the Nearest Neighbor Method. Onward, each scaled image is to be converted into points (Figure S1) for further analysis.

2.3.2. Modeling and Prediction

Regression Analysis

A polynomial regression model is established at all six upscaled resolutions (i.e., also called the low-resolution scales) using geospatial predictors to predict each individual precipitation variable. The equation deriving the relationship between geospatial predictors and precipitation variables at each low-resolution scale is as follows in Equation (4):

$$P_{D.\ GPM_{LR}} = p_1.x^3_{DEM_{LR}} + p_2.x^2_{DEM_{LR}} + p_3.x_{DEM_{LR}} + p_4 \tag{4}$$

where $P_{D.GPM_{LR}}$ is the predicted precipitation for each GPM variable at each low-resolution scale, and p_1, p_2, p_3 and p_4 are polynomial coefficients, $x^i_{DEM_{LR}}$ is the low-resolution geospatial variable(s).

In addition, a linear regression model is established to evaluate the relationship between the EDBF-based weighted precipitation and geospatial predictor, i.e., the latitude, which is as follows in Equation (5):

$$P_{D.WTP_{Res}} = p_1.x_{latitude} + p_2 \tag{5}$$

where $P_{D.WTP_{Res}}$ is the predicted weighted precipitation at investigated resolution scale(s), e.g., the low-resolution ($P_{D.WTP_{LRes}}$) or the high-resolution ($P_{D.WTP_{HRes}}$), p_1 and p_2 are linear coefficients, and $x_{latitude}$ is the geospatial predictor.

Calculation of *r* Values

To execute EDBF algorithm for predicting the weighted precipitation from the multitemporal precipitation variables, the generation of initial weight vector for each contributing precipitation variable is a key process. In this regard, the *r* value is needed to formulate the initial weight vector for

each contributing variable. The equation deriving r values at each low-resolution scale is given by Equation (6):

$$r = \frac{Cov\left(Pr_i,\ Pr_j\right)}{\left(\sigma_{Pr_i} \times \sigma_{Pr_j}\right)} \tag{6}$$

where $Cov\left(Pr_i,\ Pr_j\right)$ is the covariance and $\sigma_{Pr_i}, \sigma_{Pr_j}$ is the standard deviation of predictors (i, j) at i-th and j-th pixels, respectively. The equation deriving the covariance, the standard deviation and the mean for each investigated predictor is given by Equations (7)–(9), respectively:

$$Cov\left(Pr_i,\ Pr_j\right) = \frac{\sum_{ij=1}^{n}(Pr_i - \mu_i)\left(Pr_j - \mu_j\right)}{N} \tag{7}$$

$$\sigma = \left(\sum_{i=1}^{n}(Pr_i - \mu)^2 \Big/ N\right)^{\frac{1}{2}} \tag{8}$$

$$\mu = \frac{\sum_{i=1}^{n} Pr_i}{N} \tag{9}$$

where $Pr_i,\ Pr_j$ are the two investigated predictors; μ_i, μ_j are the mean of investigated predictors, respectively, and N is the total number of observations.

Chi-Square (χ^2) Test

This is a non-parametric test, which is used for the purpose to check the significance of relationship between each precipitation variable at each low-resolution scale for the assigned weighted values. In this regard, the assumption is made to determine the association between precipitation variables and the low-resolution scales.

The null and alternative hypothesis will be:

H_0: There is a significant relationship between precipitation variables and upscaled resolutions.
H_1: There is no significant relationship between precipitation variables and upscaled resolutions.
The equation deriving chi-square statistic [56], is given by Equation (10):

$$\chi^2 = \left(\sum_{i=1}^{n}(O_i - E_i)^2 \Big/ E_i\right) \tag{10}$$

where O_i represents the observed, and E_i represents the expected frequency.

Descriptive Statistics

Various statistical parameters are used to verify the proposed methodology by indicating the perfect score and range for each statistical metric, i.e., R^2, RMSE and the bias (B), which are expressed in Equations (11) and (12):

$$RMSE = \left(\sum_{i}^{n}(Ob_i - Pr_i)^2 \Big/ n\right)^{\frac{1}{2}} \tag{11}$$

$$B = \frac{\sum_{i=1}^{n} Pr_i}{\sum_{i=1}^{n} Ob_i} - 1 \tag{12}$$

where Ob_i is the observed variable, Pr_i is the predicted variable, and n is the number of observations.

2.3.3. EDBF Algorithm

Based on polynomial regression outputs, the most influencing geospatial predictor that predicts multitemporal precipitation variables at each low-resolution scale is considered for further evaluation through EDBF algorithm. In this research, the developed methodology is based on the earlier work of [57,58]. The execution of EDBF algorithm is shown in Figure S2. Based on calculated r values,

the process starts through randomly generating initial weight vector W, which by substituting into Equation (13) obtains WTP:

$$WTP = w_M \times M + w_{An} \times An + w_W \times W + w_{Sp} \times Sp + w_{Su} \times Su + w_{Au} \times Au + w_{Wt} \times Wt + w_{Dr} \times Dr \quad (13)$$

where WTP is the weighted precipitation, $W = \{w_M, w_{An}, w_W, w_{Sp}, w_{Su}, w_{Au}, w_{Wt}, w_{Dr}\}$ corresponds to the weight values (Equation (14)), and vector M, An, W, Sp, Su, Au, Wt and Dr corresponds to each of the eight precipitation variables, i.e., the average monthly, the average annual, the average winter, the average spring, the average summer, the average autumn, the wet year (2004) and the dry year (2001) precipitation, respectively. Additionally, vector $Res_{0.25}$, $Res_{0.75}$, $Res_{0.50}$, $Res_{1.0}$, $Res_{1.25}$ and $Res_{1.50}$ corresponds to each low-resolution scale, e.g., 0.25°, 0.5°, 0.75°, 1.0°, 1.25° and 1.50°, respectively:

$$w_M + w_{An} + w_W + w_{Sp} + w_{Su} + w_{Au} + w_{Wt} + w_{Dr} = 1 \quad (14)$$

Subsequently, the correlation coefficient $R_{WTP-Res_{0.25}}$, $R_{WTP-Res_{0.50}}$, $R_{WTP-Res_{0.75}}$, $R_{WTP-Res_{1.0}}$, $R_{WTP-Res_{1.25}}$ and $R_{WTP-Res_{1.50}}$ between the WTP and vector $Res_{0.25}$, $Res_{0.50}$, $Res_{0.75}$, $Res_{1.0}$, $Res_{1.25}$ and $Res_{1.50}$ is calculated, respectively. In addition, EDBF algorithm is run to iteratively optimize W to obtain an accurate weight vector W_t, where t represents the number of iterations. Moreover, a relationship (Equation (5)) between WTP and geospatial predictor at vector $Res_{0.25}$, $Res_{0.50}$, $Res_{1.0}$, $Res_{1.25}$ and $Res_{1.50}$ is evaluated, respectively. Hereafter, the best predicted resolution vector is used in the downscaling process. Similarly, using Equation (5), the same process is repeated for the high-resolution vector $Res_{0.05}$, i.e., 0.05° scale resolution.

3. Results

3.1. Evaluation of GPM-Based Multitemporal Precipitation

The execution of proposed downscaling methodology was first formulated through evaluating the precipitation response, e.g., each GPM-based multitemporal precipitation variable with respect to geospatial predictors at each low-resolution scale. Additionally, each investigated precipitation variable, e.g., the average monthly, the average annual (2001–2015), the average winter, the average spring, the average summer, the average autumn, the dry year (2001) and the wet year (2004) precipitation was plotted against each geospatial predictor at each upscaled resolution. Demonstration through scatter diagrams and polynomial regression (i.e., Figures S3–S5) described the relationship between precipitation variables and geospatial predictors, i.e., elevation, longitude and latitude at upscaled resolutions, respectively. Moreover, the R^2 values are shown in Table 1, wherein all precipitation variables showed strong response to latitude followed by longitude and elevation, respectively. Furthermore, for the individual precipitation variables, the average spring followed by the dry year (2001) and the wet year (2004) precipitation showed a higher relationship with geospatial predictors, respectively. Apart from geospatial predictors, the highest R^2 for upscaled resolutions was observed at 1.0° and 0.75°, respectively.

Table 1. Output of model fitting between the multitemporal GPM variables and geospatial predictors.

SR	GP	Multitemporal Precipitation							
		M	A	Wn	Sp	Su	Au	Wet-y	Dry-y
	Elevation	0.0804	0.0805	0.0614	0.1636	0.0171	0.0407	0.0567	0.1052
	Longitude	0.0641	0.0636	0.209	0.1473	0.0618	0.1792	0.0476	0.0768
0.25°	Latitude	0.6692	0.6684	0.512	0.7974	0.5793	0.5262	0.7416	0.7672
	Elevation	0.0906	0.0906	0.0573	0.1763	0.0166	0.0452	0.0686	0.1033
	Longitude	0.0527	0.0527	0.1883	0.1314	0.0563	0.1855	0.0414	0.0716
0.50°	Latitude	0.5658	0.6558	0.5023	0.786	0.5913	0.5188	0.7422	0.7648
	Elevation	0.0516	0.0906	0.0516	0.1346	0.0069	0.0372	0.0343	0.0875
	Longitude	0.05	0.05	0.205	0.1488	0.0377	0.1891	0.0358	0.0762
0.75°	Latitude	0.6853	0.6853	0.5271	0.8097	0.6148	0.5432	0.7413	0.7848
	Elevation	0.0579	0.0579	0.0517	0.1534	0.0128	0.041	0.0357	0.1066
	Longitude	0.0722	0.0722	0.2006	0.1328	0.0825	0.1329	0.0686	0.0767
1.0°	Latitude	0.6929	0.6929	0.5184	0.8197	0.6415	0.5206	0.7837	0.7848
	Elevation	0.152	0.152	0.605	0.2963	0.0176	0.0257	0.1025	0.2255
	Longitude	0.0554	0.0554	0.2227	0.1394	0.0708	0.2524	0.0369	0.0656
1.25°	Latitude	0.6824	0.6824	0.532	0.8239	0.5685	0.493	0.7814	0.7802
	Elevation	0.2056	0.2056	0.207	0.2507	0.0646	0.1363	0.1543	0.1867
	Longitude	0.053	0.053	0.2133	0.1143	0.0694	0.24	0.0405	0.0682
1.50°	Latitude	0.6179	0.6179	0.4778	0.7699	0.5392	0.5325	0.6889	0.7491

Note: SR stands for the scaled resolution, GP. for geospatial predictors, M for the average monthly, A for the average annual (2001–2015), Wn for the average winter, Sp for the average spring, Su for the average summer, Au for the average autumn, Wet-y for the wet year (2004), Dry-y for the dry year (2001) precipitation, respectively.

3.2. EDBF-Based Weighted Precipitation

3.2.1. Low-Resolution Weighted Precipitation

Based on initial regression analysis, the most influencing geospatial predictor, namely latitude, was selected to predict the weighted precipitation from the multitemporal precipitation variables via EDBF algorithm. In this regard, *r* values were calculated (Figure 3a), and used as the basis function to randomly assign initial weight value to each precipitation variable. The reason for negative *r* values is the existence of a negative relationship between latitude and precipitation variables. Subsequently, the chi-square (χ^2) test was employed to evaluate the relationship between precipitation variables at each low-resolution scale for assigned weight values. The $\chi^2_{calculated}$ and $\chi^2_{tabulated}$ values with 35 degrees of freedom at the significance level ($\alpha = 0.05$) were 5.267 and 49.802, respectively. Based on statistical results, the $\chi^2_{calculated} < \chi^2_{tabulated}$, thus, the null hypothesis was accepted and rejected the alternative hypothesis. Moreover, it is stated that weight values assigned to precipitation variables were significantly not different. The details can be found in Table S1.

Furthermore, the correlation between precipitation variables and the low-resolution scales was analyzed and is shown in Figure 3b, wherein it showed that the dry year (2001) followed by the average spring, the wet year (2004) and the average summer precipitation are the most influencing variables. As far the scaled resolutions are concerned, 1.0° followed by 0.75° resolution had higher impacts.

Onward, the precipitation data was evaluated through EDBF algorithm, and the number of iterations was set to 3×10^4. Figure 4 demonstrates the iteration wise statistics at each upscaled resolution, in which Figure 4a,d,g,j,m,p show weight values, Figure 4b,e,h,k,n,q show *r* values, and Figure 4c,f,i,l,o,r show the comparison between weight and *r* values which were iteratively generated by the algorithm itself. To investigate weight values, it was observed that lots of discrepancies exist in the convergence of investigated variables (e.g., Figure 4a,d,g,j,m,p), and the convergence showed stabilization onward 2×10^4 iterations. In Figure 4a,d,g, the dry year (2001), the average spring and the average autumn; Figure 4j, the dry year (2001), the average spring and the wet year (2004); Figure 4m,p, the dry year (2001), the average autumn and the average spring, respectively, showed higher weight values from the beginning until the last iteration. As far *r* values are concerned, uncertainty in initial iterations was observed as shown in Figure 4b,e,h,k,n,q, and the convergence

showed stabilization onward 1×10^4 iterations. Likewise, it was also observed that *r* values drastically decreased before the stabilization of convergence.
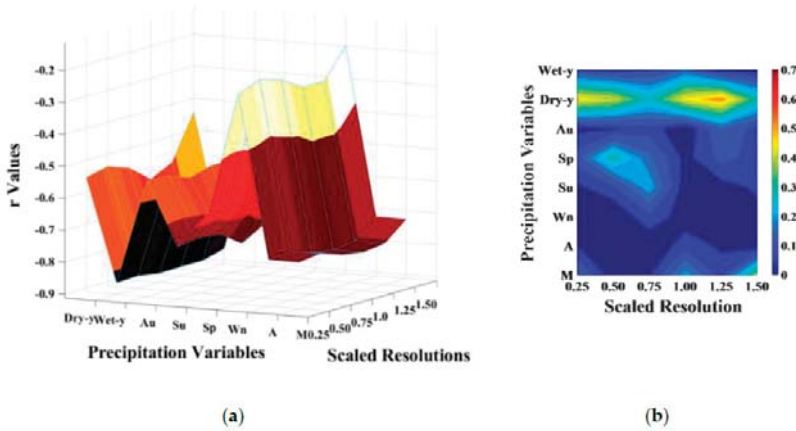
Figure 3. Selection of initial input parameters for EDBF) algorithm to predict the weighted precipitation (**a**) calculated *r* values, and (**b**) the correlation between precipitation variables and the low-resolution scales for assigned weight values.

In addition, the weighted *r* value predicted by EDBF algorithm was higher as compared to the calculated *r* value for each precipitation variable, as shown in Figure 4c,f,i,l,o,r. The highest weighted *r* was predicted at 1.0° (−0.891) followed by 0.75° and 1.25° (−0.889), 0.25° (−0.880), and 0.50° and 1.50° (−0.867), respectively. Nevertheless, the final weight value predicted at all upscaled resolutions was same, e.g., equal to 1, but weighted response towards precipitation variables was not similar. It can clearly be observed that the highest weighted response was given to the dry year (2001) (Figure 4l,i,c,f,r,o) followed by the average spring (Figure 4l,i,o,c,f,r), the average autumn (Figure 4f,r,o,c,i), and the wet year (2004) (Figure 4l), respectively. Finally, the relationship between latitude and the weighted precipitation predicted by EDBF algorithm was shown through scatter plots in Figure 5a–f. In contrast to earlier plots, i.e., the exitance of polynomial relationship between precipitation variables and geospatial predictors, here, the linear relationship was observed. Moreover, the R^2 between latitude and the weighted precipitation at each upscaled resolution was increased. The higher R^2 was observed at 1.0°, 1.25°,0.75° resolutions, respectively. Overall, R^2 was higher than 0.75.
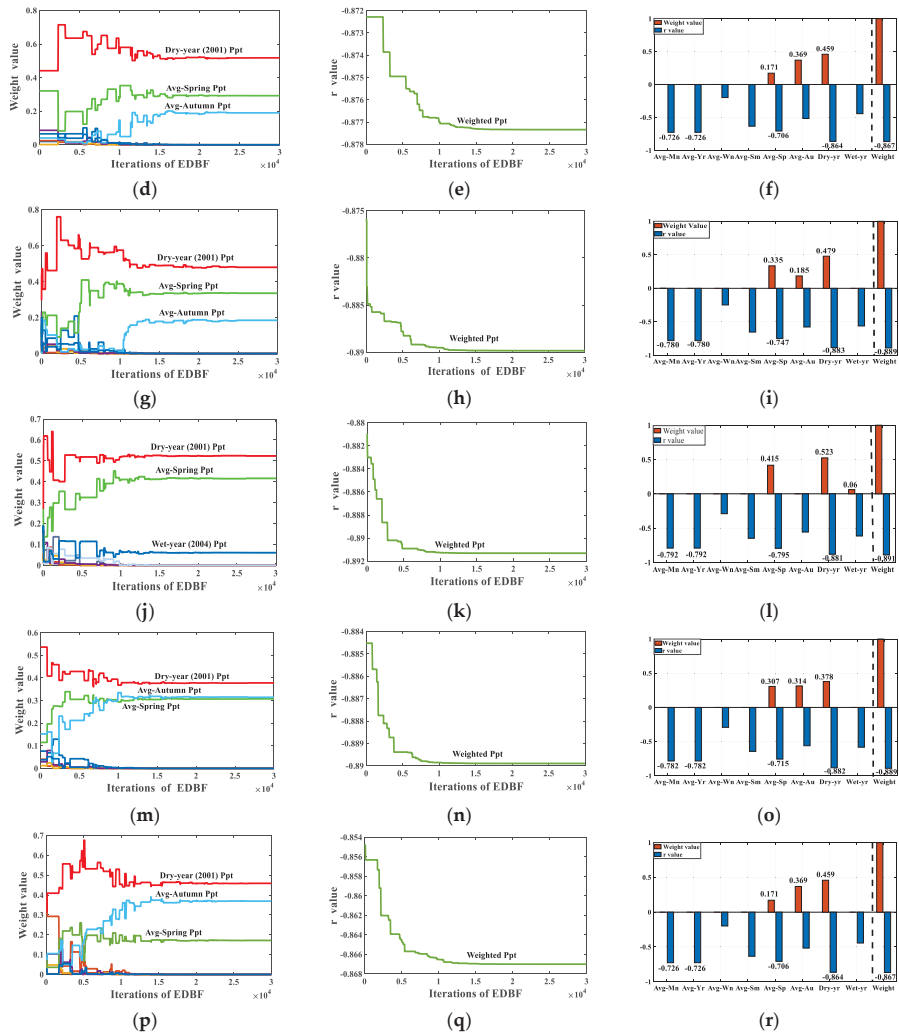
Figure 4. Execution of EDBF algorithm at different low-resolution scales, (**a**,**d**,**g**,**j**,**m**,**p**) iteratively estimated weighted values, (**b**,**e**,**h**,**k**,**n**,**q**) iteratively estimated *r* values, and (**c**,**f**,**i**,**l**,**o**,**r**) the comparison between assigned weights and estimated *r* values at 0.25°, 0.50°, 0.75°, 1.0°, 1.25°, 1.50° resolutions, respectively.
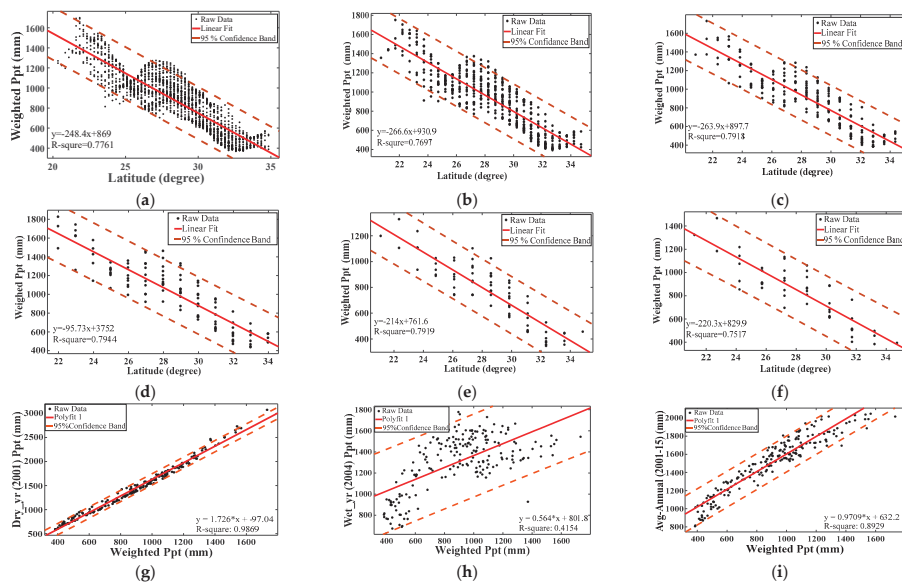
Figure 5. Scaled wise relationship between the weighted precipitation and (**a**) latitude at 0.25° resolution; (**b**) latitude at 0.50° resolution; (**c**) latitude at 0.75° resolution; (**d**) latitude at 1.0° resolution; (**e**) latitude at 1.25° resolution; (**f**) latitude at 1.50° resolution; (**g**) the dry year (2001) precipitation at 0.75° resolution; (**h**) the wet year (2006) precipitation at 0.75°; and (**i**) the average annual (2001–2015) precipitation at 0.75°, respectively.

3.2.2. High-Resolution Weighted Precipitation

To predict the weighted precipitation at 0.05° resolution by EDBF algorithm, the same process was adopted as done for the low resolutions. In contrast to predict the low-resolution weighted precipitation, here, the number of iterations was reduced and set to 1000. Figure 6 shows the iteration wise statistics, wherein Figure 6a,b, respectively, show the iteration wise weight values for each precipitation variable and *r* values. In both figures, the initial discrepancies were observed before the stabilization of convergence onward 600 iterations. On top of that, during the prediction of weighted precipitation, the higher weighted response (Figure 6a,d) was shown by the dry year (2001), the average spring and the average autumn precipitation, respectively. The true picture of weight given to each precipitation variable based on *r* value was cleared from Figure 6c, wherein it can clearly be observed that the highest weighted response was given to the dry year (2001) ($w_{Dr} = 0.507$) followed by the average spring ($w_{Sp} = 0.220$), the average autumn ($w_{Au} = 0.151$), the average monthly ($w_M = 0.050$), the average summer ($w_{Su} = 0.046$), the average winter ($w_W = 0.015$), the wet year (2004) ($w_{Wt} = 0.005$) and the average annual (2001–2015) precipitation ($w_{An} = 0.002$), respectively. Moreover, the comparison between weight and *r* values are shown in Figure 6d, wherein it showed that the weighted *r* value predicted by EDBF algorithm is higher as compared to the calculated *r* value for each precipitation variable, which reflected that the weighted precipitation showed more consistency as compared to individual precipitation variable. Finally, the relationship between the weighted precipitation predicted by EDBF algorithm and latitude (Figure 6e) was shown through scatter plot, and the achieved R^2 was observed 0.7696.
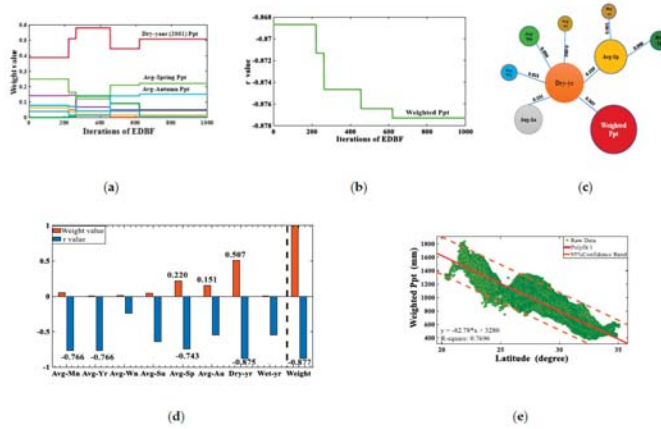
Figure 6. Execution of EBDF algorithm for predicting the high-resolution weighted precipitation at 0.05° resolution: (**a**) iteratively estimated weight values; (**b**) iteratively estimated *r* values; (**c**) the estimated final weight values; (**d**) the comparison between the estimated weights and r values; and (**e**) the relationship between the weighted precipitation and latitude, respectively.

3.3. Verification Process

3.3.1. Comparison between the Weighted and the Original Multitemporal Precipitation Variables

To compare the EDBF-based weighted precipitation with the GPM-based multitemporal precipitation variables, a linear regression model was established at all upscaled resolutions, e.g., 0.25°, 0.5°, 0.75°, 1.0°, 1.25° and 1.50°. The efficiency comparison was established using three statistical metrics, i.e., R^2, RMSE, and the bias (B). The results are shown in Table 2 (e.g., 0.75° resolution) and Table S2 (0.25°, 0.5°, 1.0°, 1.25° and 1.50°), respectively. From the tabulated results, it was observed that for the achieved R^2 value, the weighted precipitation outperformed all multitemporal variables at all upscaled resolutions. The highest R^2 value of 0.794 was observed at 1.0° followed by 0.792 at 0.75° resolution, respectively. Also, for the achieved RMSE value, the weighted precipitation outperformed the annual precipitation variables, such as the average annual (2001–2015), the wet year (2004) and the dry year (2001) precipitation, whereas it underperformed compared to the seasonal, e.g., the average winter, the average spring, the average summer and the average autumn precipitation, and the monthly precipitation variables. The lowest RMSE value (i.e., at all upscaled resolutions) was observed for the average monthly precipitation. Moreover, the observed bias for the two precipitation datasets, e.g., the weighted precipitation and the multitemporal precipitation variables, was almost reaching zero. In addition, both precipitation datasets were also compared at the original 0.1° resolution as shown in Table 3. The tabulated results revealed the same outcome as in Table 2, wherein the best correlation (R^2) was observed between latitude and the weighted precipitation, and it outperformed all multitemporal precipitation variables. Similarly, for the achieved RMSE value, it outperformed the annual precipitation variables and underperformed compared to the seasonal and the monthly variables. As a whole, the observed output at each statistical parameter for each precipitation variable was slightly reduced from lower to higher (e.g., from Table 2 to Table 3) resolution.

Table 2. Comparison between the weighted precipitation and the multitemporal precipitation variables at 0.75° resolution.

Variables	Statistical Parameters				
	R-Square	Mean	S. D.	RMSE	Bias
Weighted Ppt	0.792	897.680	295.860	135.005	−3.0E−06
Avg-Monthly	0.609	125.319	25.335	15.836	−3.6E−05
Avg-Annual	0.609	1503.828	304.016	190.043	2.8E−05
Avg-Winter	0.533	181.795	70.273	48.041	4.5E−05
Avg-Summer	0.426	617.508	123.003	93.165	8.1E−06
Avg-Spring	0.559	466.215	160.574	106.693	−0.00012
Avg-Autumn	0.333	241.951	47.712	38.959	−2.2E−05
Wet-Y (2004)	0.316	1308.082	258.892	214.166	1.9E−05
Dry-Y (2001)	0.781	1452.542	514.112	240.376	0.00010

Table 3. Comparison between the weighted precipitation and multitemporal precipitation variables original 0.1° resolution.

Variables	Statistical Parameters				
	R-Square	Mean	S. D.	RMSE	Bias
Weighted Ppt	0.772	920.915	295.731	141.113	8.89E−05
Avg-Monthly	0.591	124.681	25.084	16.044	2.17E−05
Avg-Annual	0.591	1496.164	301.015	192.537	3.3E−05
Avg-Winter	0.510	181.066	70.465	49.303	−0.00034
Avg-Summer	0.407	612.638	119.603	92.099	−5.1E−06
Avg-Spring	0.561	464.828	159.760	105.901	−5.2E−05
Avg-Autumn	0.299	241.260	49.556	38.979	3.78E−07
Wet-Y (2004)	0.307	1306.001	255.948	212.987	−3.2E−05
Dry-Y (2001)	0.766	1440.258	503.257	243.108	−5.7E−05

3.3.2. Verification of the Weighted Precipitation with Neutral Variables

The weighted precipitation was further evaluated by comparing with neutral variables which were not used during the prediction of EDBF-based weighted precipitation. In this regard, the precipitation variables from two different datasets, such as the TRMM and the GPM, were used for the verification of weighted precipitation. The GPM dataset used during verification comprised of the annual 2006 (Figure 2j) and 2012 (Figure 2k) precipitation, whereas the TRMM dataset comprised of the annual 2001, 2006, 2012 and the average annual (2001–2015) (Figure 2l–o) precipitation, respectively. The verification of weighted precipitation though the GPM data was evaluated by extracting precipitation at the original 0.1° resolution, whereas through the TRMM data, it was evaluated at the original 0.25° resolution. The verification results are shown in Table 4. The weighted precipitation outperformed both, as can be observed by comparing the datasets, by achieving a higher R^2 value of 0.776 at 0.25° resolution and 0.772 at 0.1° resolution as compared to the TRMM and the GPM-based precipitation, respectively. Subsequently, the weighted precipitation also produced lower RMSE, e.g., 133.37 (0.25° resolution) and 141.113 (0.1° resolution) as compared to the TRMM- and the GPM-based precipitation, respectively. Apart from that, the observed bias almost reached zero for all variables, wherein the weighted precipitation showed positive bias, while the TRMM and the GPM precipitation showed negative bias.

Table 4. Comparison between the weighted precipitation and neutral precipitation variables.

Variables	Statistical Parameters				
	R-Square	Mean	S. D.	RMSE	Bias
Weighted Ppt @ 0.1°	0.772	920.915	295.731	141.113	8.8E−05
GPM (2006)	0.668	1555.048	463.560	267.117	−0.00013
GPM (2012)	0.391	1687.073	517.823	404.087	4.3E−05
Weighted Ppt @ 0.25°	0.776	868.975	281.890	133.377	9.5E−06
TRMM (2001)	0.599	1393.780	480.101	303.951	−1.8E−05
TRMM (2012)	0.318	1668.233	546.886	451.363	−1.1E−05
TRMM (2006)	0.473	1608.031	542.391	393.678	−2.1E−05
Avg-TRMM (01-15)	0.373	1534.130	374.400	296.581	−2.7E−05

3.4. Downscaling of the Weighted Precipitation

Based on the verification of EDBF results, the algorithm was employed in the downscaling process. During the downscaling process, a distinction between the low-resolution (upscaling) and the high-resolution (downscaling) was made by using Equation (5). By subtracting the weighted precipitation $P_{D.WTP_{LR}}$ (i.e., also called the low-resolution weighted precipitation) (Figure 7b) from the original Avg_MT_{GPM} precipitation (Figure 7a), the residuals $\left(R_{WTP_{LR}}\right)$ of the regression model (i.e., also called as the low-resolution weighted residuals) at 0.75° resolution were obtained as shown in Figure 7c, which represents the amount of precipitation that could not be predicted by the weighted precipitation via EDBF algorithm according to Equation (15). Subsequently, the generated residuals were interpolated to 0.05° resolution (Figure 7d), also called the high-resolution weighted residuals ($R_{WTP_{HR}}$), by applying a spline tension interpolator [59]. Finally, the high-resolution weighted precipitation ($P_{D.WTP_{HR}}$) at 0.05° resolution (Figure 7e) was obtained using Equation (5). Using Equation (16) to add the high-resolution weighted precipitation to the high-resolution weighted residuals, the final downscaled high-resolution weighted precipitation $P_{Ds.PWTP_{HR}}$ (Figure 7f) for the humid region of Mainland China was obtained:

$$R_{WTP_{LR}} = Avg_{GPM} - P_{D.WTP_{LR}} \tag{15}$$

$$P_{Ds.PWTP_{HR}} = P_{D.WTP_{HR}} + R_{WTP_{HR}} \tag{16}$$
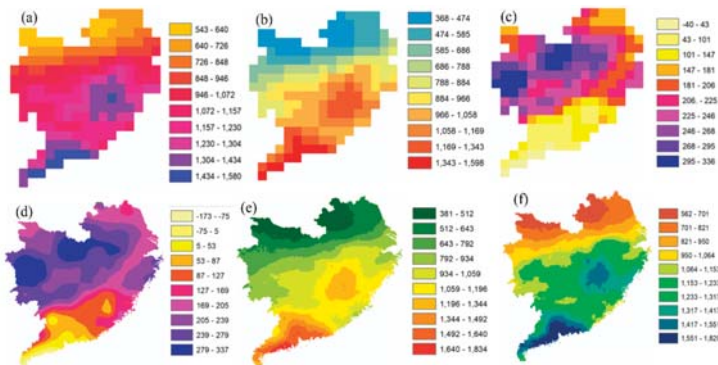
Figure 7. Stepwise downscaling process to predict the high-resolution multitemporal weighted precipitation: (**a**) the GPM-based average multitemporal precipitation at 0.75° resolution; (**b**) the EDBF-based weighted precipitation at 0.75° resolution; (**c**) the low-resolution weighted residuals at 0.75° resolution generated from the difference between (**a**) and (**b**); (**d**) the high-resolution weighted residuals at 0.05° resolution generated by interpolating (**c**) via Spline Interpolation; (**e**) the EDBF-based high-resolution weighted precipitation at 0.05° resolution; and (**f**) the final high-resolution downscaled multitemporal weighted precipitation, at 0.05° resolution, as a product of adding (**d**) into (**e**), respectively.

4. Discussion

In this study, a new downscaling methodology, namely GMWPA at 0.05° resolution, was developed and investigated in the humid region of Mainland China. A two-stepped procedure [38,39,41], based on a scale-dependent regression analysis and downscaling of the predicted multitemporal weighted precipitation at a refined scale, was adopted during the execution of proposed methodology. For this purpose, the multitemporal GPM precipitation dataset (2001 to 2015) at 0.1° and ASTER 30 m DEM-based geospatial predictors, i.e., elevation, longitude, and latitude were taken as input variables to predict the low-resolution—for the residual generation at optimal resolution scale—and the high-resolution weighted precipitation, and were used in the final downscaling process.

Furthermore, the regression analysis was performed in two phases. In the first phase, each geospatial predicator was assessed through developing a relationship (Table 1) with each individual precipitation variable via a fitting line—polynomial fit. Moreover, it was observed that latitude showed the highest correlation with all precipitation variables and achieved the highest R^2 value. Compared to previous studies [3,34,59] which used either one or two independent variables (NDVI, elevation), the authors in [38] used several independent variables, i.e., latitude, longitude, elevation, slope, aspect, NDVI, Max_NDVI, Range_NDVI, and Min_NDVI, to establish regression models for deriving the annual precipitation over continental China. From the study, it was concluded that, apart from latitude, all variables including NDVI showed relatively weak empirical relationships with the observed precipitation, especially over the humid region of China. Specifically, for NDVI, a possible reason may be that NDVI-related predictors are better indicator of precipitation in arid and semi-arid areas. The NDVI values would not increase with the increased rainfall amount in humid areas, which makes a relatively weak empirical relationship between precipitation and saturated NDVI. Keeping in view, latitude was selected as the proxy of precipitation and employed in assigning initial weight value (e.g., based on *r* value calculated for each precipitation variable with respect to latitude) to each individual precipitation variable from the multitemporal precipitation dataset, and which was then processed in EDBF algorithm [58] to predict the weighted precipitation.

Likewise, in the second phase, the output precipitation variable from EDBF, e.g., the weighted precipitation was assessed via developing the relationship with latitude through linear fitting. Moreover, the correlation between latitude and the weighted precipitation was increased for each of the low-resolution scale, and the highest R^2 was achieved at 100 km (e.g., between 0.75°, 1.0°, 1.25° resolutions), which showed that the weighted precipitation was well captured by latitude at 100 km resolution. Although the highest correlation between latitude and the weighted precipitation was achieved at 1.0° (100 km), but due to certain reasons, 0.75° resolution was selected as an optimal low resolution (e.g., for the upscaling) during the downscaling process. First, there was not much difference between the two resolution scales for the achieved R^2, i.e., 0.75° ($R^2 = 0.7918$) and 1.0° ($R^2 = 0.7977$) resolution. Secondly, 0.75° resolution had more pixels, i.e., 195, as compared to 111 pixels for 1.0° resolution to cover the whole study area. Considering, to convert points into pixels, the Spline Interpolation method [51,60] was used, which estimates values using a mathematical function that minimizes the overall surface curvature, resulting in a smooth surface that passes exactly through a specified number of nearest input points while passing through the sample points. Thus, using 0.75° resolution, which had a closer specified number of nearest input points, i.e., 12 points, than 1.0° resolution, tends to produce a smoother surface by minimizing the surface curvature.

From the EDBF algorithm perspective, it is a general framework rather than a specific algorithm, which is easy to implement and can easily accommodate any existing multi-parent crossover algorithms (MCAs). Moreover, the existing MCA-based coefficients [61–63] follow a uniform distribution, which also violates constraints, thus propagate error. Errors cascade exponentially, with even a slight increase in the hybrid scale, which leads to the increase in time consumption. To address such problem, EDBF is the best solution which takes multiple MCAs as its constituent members. In addition, the number of iterations during the execution of EDBF algorithm at the low-resolution scale, i.e., 0.25°, 0.50°, 0.75°, 1.0°, 1.25° and 1.50° was set to 3×10^4 with the reason that a possible number of iterations

be available for the stabilization of convergence before the ending of simulation process. Moreover, the process was repeated for all the low resolutions. Though the convergence stabilized before a 3×10^4 number of iterations, still a slight improvement could be observed, and further improvement in the regression value(s) could be expected. Instead, by terminating simulation during the execution, we let simulation process to be completed until the last iteration. Owing to that, the number of iterations was reduced during the simulation of high-resolution (i.e., 0.05° resolution) weighted precipitation, and the convergence was well stabilized within the set number of iterations.

During the verification process, the weighted precipitation was first compared with its contributing multitemporal precipitation variables at all the low and the original resolution scales. It outperformed all input variables for the achieved R^2 and outperformed the annual precipitation and underperformed compared to the seasonal and the monthly precipitation variables for the achieved RMSE. Furthermore, the weighted precipitation was compared with different classified precipitations, extracted either as an individual or grouped variables from the original multitemporal precipitation dataset used in the prediction of EDBF-based weighted precipitation at the original 0.1° resolution. The results are shown in Table 5, in which the weighted precipitation showed the highest correlation with its predictor (R^2 = 0.772) as compared to other used variables. In addition, the weighted precipitation had a lower RMSE value (e.g., RMSE = 141.113 mm) than the Avg-An (01–15) + Wet Ppt+ Dry Ppt, Avg-An (01–15) + Dry Ppt, Avg-An (01–15) + Wet Ppt, Wet Ppt + Dry Ppt, Avg-An (01–15) and Avg-MT (−01 & −04) Ppt with the observed RMSE value of 179.248, 206.353, 182.762, 178.025, 192.537 and 197.434 mm, respectively. Also, it had a higher RMSE than the Avg-MT Ppt variable, i.e., 135.370 mm. The reason of low RMSE value for the average multitemporal GPM precipitation was that the average output was equally contributed by each precipitation variable from the multitemporal dataset. Out of the eight used variables from the multitemporal precipitation dataset, the five variables consisted of the seasonal and the monthly precipitation, which had lower received pixel precipitation. Adding to this, the number of days counted during each of the seasonal component (e.g., average 90 days) is lower than the annual component (e.g., 365 days) and there is less probability of variation in the seasonal precipitation than the annual precipitation. Despite lower R^2 values, less variability from the mean precipitation was observed in the seasonal and the monthly precipitation as compared to the annual precipitation. On the contrary, the EDBF-based weighted precipitation was mainly predicted on the basis of assigned weights via calculated *r* values. In this regard, higher the *r* value, the more weight was assigned to that variable and more contribution from that variable in the prediction of weighted precipitation. Additionally, it was compared with neutral variables, wherein it outperformed all comparing variables for the achieved R^2 and RMSE values.

Table 5. Comparison between the weighted precipitation and classified extracted precipitation variables.

Variables	Statistical Parameters				
	R-Square	Mean	S. D.	RMSE	Bias
Weighted Ppt	0.772	920.915	295.731	141.113	8.8E−05
Avg-MT Ppt	0.708	1096.535	250.712	135.370	5.8E−06
Avg-An (01-15) + Wet + Dry Ppt	0.696	1414.141	325.185	179.248	−2.1E−05
Avg-An (01-15) + Dry Ppt	0.726	1468.211	394.319	206.353	−9.4E−05
Avg-An (01-15) + Wet Ppt	0.511	1401.083	261.236	182.762	1.2E−05
Wet + Dry Ppt	0.728	1373.130	341.383	178.025	1.2E−05
Avg-An (01-15)	0.591	1496.164	301.015	192.537	3.3E−05
Avg-MT (-01 & -04) Ppt	0.558	1515.092	297.221	197.434	−2.9E−05

Note: Avg-MT Ppt is the average multitemporal precipitation; Avg-An(01-15) + Wet + Dry Ppt is the average precipitation as product of the average annual, the wet year (2004) and the dry year (2001) precipitation; Avg-An + Dry Ppt is the average precipitation as product of the average annual (2001–2015) and the dry year (2001) precipitation; Avg-An(01-15) + Wet Ppt is the average precipitation as product of the average annual (2001–2015) and the wet year (2004) precipitation; Wet + Dry Ppt is the average precipitation as product of the wet year (2004) and the dry year (2001) precipitation; Avg-An (01-15) is the average annual (2001–2015) precipitation; Avg-MT(-01 & -04) Ppt is the average multitemporal precipitation excluding the dry year (2001) and the wet year (2004) precipitations.

The downscaling methodology applied in this study was mainly based on the work presented in [39], where the basis function was selected at an optimum resolution and by interpolating the residuals. After successfully applying the proposed methodology, the EDBF algorithm was employed in downscaling of the dry year (2001), the wet year (2004) and the average annual (2001–2015) precipitation at 0.05° resolution by following the same process as for downscaling the Avg_MT_{GPM} precipitaiton. Before downscaling, a graphical relationship between the weighted precipitiaon and the dry year (2001), the wet year (2004) and the average annual (2001–2015) precipitation was developed through a scatter plot as shown in Figure 5g–i, respectively. The weighted precipitation showed the highest correlation with the dry year (2001) followed by the average annual (2001–2015) and the wet year (2004) for the achieved R^2 = 0.9869, 0.8929 and 0.4154, respectively.

Moreover, during downscaling, the low-resolution weighted residuals (Figure S6d–f) were generated by subtracting the low-resolution weighted precipitation $P_{D.WTP_{LR}}$ (Figure 7b) from the original dry year (2001), the wet year (2004) and the average annual (2001–2015) precipitation (Figure S6a–c) at 0.75° resolution, respectively. Afterward, the high-resolution weighted residuals (Figure S6g–i) at 0.05° were obtained by interpolating the low-resolution residuals at 0.75° resolution. Finally, by adding the obtained high-resolution interpolated residuals to the high-resolution weighted precipitation (Figure 7e), the downscaled high-resolution weighted precipitation at 0.05° resolution for the dry year (2001) (Figure 8d), the wet year (2004) (Figure 8e) and the average annual (2001–2015) precipitation (Figure 8f) was obtained. From Figure 8, it shows that the high-resolution weighted precipitation captured the same precipitation pattern as that of the original GPM dry year (2001), the wet year (2004) and the average annual (2001–2015) precipitation at 0.1°. Moreover, by analyzing the class wise pattern (Table 6) for the obtained precipitation, the algorithm accurately captured the wet year (2004) (Figure 8e) and the average annual (2001–2015) precipitation, whereas some classes, e.g., class 4 (gold color) and 5 (light green) were not very well captured during downscaling of the dry year (2001) precipitation, such as between 111° to 115°E and 25° to 27°N, and 117° to 118°E and 24° to 25°N.
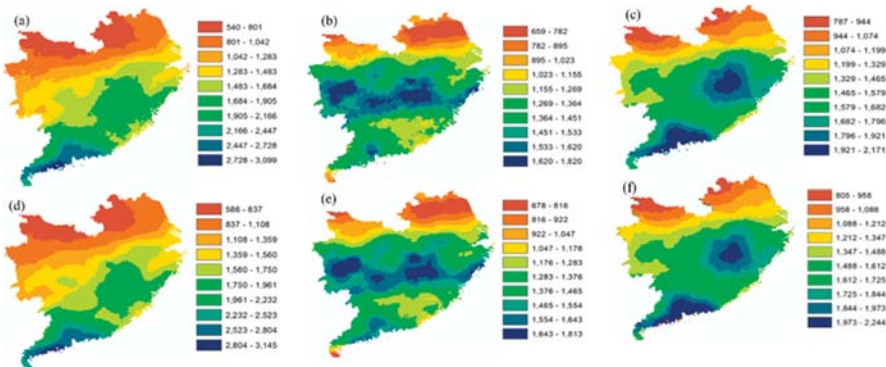
Figure 8. Comparison between the downscaled weighted precipitation at 0.05° resolution for (**d**–**f**) and the original GPM precipitation at a nominal resolution of 0.1° for (**a**) the dry year (2001), (**b**) the wet year (2004), and (**c**) the average annual (2001–2015), respectively.

Table 6. Comparison of spatial pattern between GPM based precipitations with its corresponding weighted precipitation for different precipitation classes.

Variables	Precipitation Classes									
	540	801	1042	1283	1483	1684	1905	2166	2447	2728
	801	1042	1283	1483	1684	1905	2166	2447	2728	3099
GPM 2001	261	241	241	200	201	221	261	281	281	371
	586	837	1108	1359	1560	1750	1961	2232	2523	2804
	837	1108	1359	1560	1750	1961	2232	2523	2804	3145
Weighted 2001	251	271	251	201	190	211	271	291	281	341
	659	782	895	1023	1155	1269	1364	1451	1533	1620
	782	895	1023	1155	1269	1364	1451	1533	1620	1820
GPM 2004	123	113	128	132	114	95	87	82	87	200
	676	816	922	1047	1176	1283	1376	1465	1554	1643
	816	922	1047	1176	1283	1376	1465	1554	1643	1813
Weighted 2004	140	106	125	129	107	93	89	89	89	170
	787	944	1074	1199	1329	1465	1579	1682	1796	1921
	944	1074	1199	1329	1465	1579	1682	1796	1921	2171
Avg-Annual (2001–2015)	157	130	125	130	136	114	103	114	125	250
	805	958	1088	1212	1347	1488	1612	1725	1844	1973
	958	1088	1212	1347	1488	1612	1725	1844	1973	2244
Weighted Avg-Annual (2001–2015)	153	130	124	135	141	124	113	119	129	271

Note: Green colored values show the average difference of less than 5 mm, Blue colored values show the average difference of about 10 mm, Pink colored values show the average difference of about 20 mm, Red colored values show the average difference of about 30 mm between the GMP and weighted precipitation at particular pattern class, respectively.

Subsequently, to analyze difference in the range of precipitation classes (i.e., difference between the upper and the lower boundary of captured precipitation pattern) between the original dry year (2001), the wet year (2004) and the average annual (2001–2015) precipitation at 0.1° resolution, their corresponding weighted precipitation at 0.05° resolution was found to be in close proximity with the average difference of less than 5 mm for most classes. Apart from that, EDBF algorithm slightly underpredicted extreme precipitation for the dry year (2001) and the wet year (2004) with the average difference of 30 mm, and overpredicted the average annual (2001–2015) precipitation with a difference of 20 mm. On the contrary, for low precipitation EDBF underpredicted the dry year (2001) and the average annual (2001–2015), and overpredicted the wet year (2004) precipitation with the average difference of 10, 4 and 23 mm, respectively. Similarly, considering the individual precipitation variable, EDBF accurately predicted the wet year (2003) and the average annual (2001–2015) precipitation with the average difference of less than 5 mm, whereas it slightly overpredicted the dry year (2001) with an average difference of 10 mm between the original and the corresponding weighted precipitation.

5. Conclusions

This study investigated and developed a new downscaling methodology, such as GMWPA at 0.05° resolution based on the multitemporal GPM precipitation dataset (2001 to 2015) at 0.1° and ASTER 30 m DEM-based geospatial predictors, i.e., elevation, longitude, and latitude in EDBF algorithm. The proposed methodology is a two-stepped process: (i) to develop a scale dependent relationship between precipitation variables, i.e., the multitemporal GPM precipitation and the weighted precipitation, and geospatial predictors through regression analysis [45]; (ii) the downscaling of EDBF-based multitemporal weighted precipitation at a refined scale. In addition, EDBF results were validated using neutral variables, e.g., the GPM-based annual 2006 and 2012 precipitation, the TRMM-based annual (2001, 2006 and 2012) and the average annual (2001–2015) precipitation. The following conclusions are drawn from this work:

- Geospatial predictors were the proxy of precipitation and polynomial function best described the relationship between the multitemporal precipitation variables and geospatial predictors, i.e., elevation, longitude, and latitude.

- The correlation between the multitemporal GPM variables and geospatial predictors varies with resolution, and the best correlation was found at a resolution of approximately 100 km (0.75°–1.25°). The highest correlation between precipitation variables and geospatial predictors was observed for the average spring followed by the dry year (2001) and the wet year (2004) precipitation, respectively. The latitude showed to be the best geospatial predictor.

- The weighted *r* value predicted by EDBF algorithm was higher than the calculated *r* value for each of the individual precipitation variables. The highest weighted *r* value was predicted at 1.0° (−0.891) followed by 0.75°and 1.25° (−0.889), respectively. Besides, the highest weighted response was observed for the dry year (2001), followed by the average spring, the average autumn and the wet year (2004), respectively.

- In contrast to the priori polynomial relationship between the multitemporal precipitation variables and geospatial predictors, a consistent linear relationship between the weighted precipitation and latitude was observed with an R^2 value of 0.7696, 0.7761, 0.7697, 0.7918, 0.7944, 0.7919 and 0.7517 at 0.05°, 0.25°, 0.50°, 0.75°, 1.0°, 1.25° and 1. 50° resolution, respectively.

- In comparison with the multitemporal GPM variables, the weighted precipitation outperformed all variables for the achieved R^2 value, whereas it outperformed the annual precipitation variables and underperformed compared to the seasonal and the monthly variables for the achieved RMSE value. In addition, it outperformed all comparing variables during the verification process for the achieved R^2 and RMSE values.

- Based on achieved results, the downscaling process was carried out for the average multitemporal precipitation, the multitemporal annual precipitation (2001 and 2004) and the average annual precipitation (2001–2015).

- The proposed downscaling methodology was refined through earlier methodologies described in [3,39,64] by selecting the basis function at an optimum resolution and by interpolating the residuals.

- The downscaling approach resulted through the proposed methodology captured spatial patterns with greater accuracy at higher spatial resolution.

- This work showed that it is feasible to increase the spatial resolution and accuracy of a precipitation variable on an annual basis or as an average from the multitemporal precipitation dataset using a geospatial predictor, i.e., latitude as the proxy of precipitation through the weighted precipitation. Future work should focus on extending this procedure using the multitemporal precipitation dataset from multi-satellites or a satellite combining rain gauge precipitation, also through analyzing the combined effect of predictors (e.g., geospatial and environmental, etc.) as the proxy of precipitation.

In conclusion, it is possible to accurately downscale the GPM-based multitemporal precipitation using geospatial predictors in the humid region (Southern China) of Mainland China and that the presented methodology is generic in nature and is applicable in all climatic conditions of the world.

References

1. Langella, G.; Basile, A.; Bonfante, A.; Terribile, F. High-resolution space–time rainfall analysis using integrated ANN inference systems. *J. Hydrol.* **2010**, *387*, 328–342. [CrossRef]

2. Lopez, P.; Immerzeel, W.; Rodríguez Sandoval, E.; Sterk, G.; Schellekens, J. Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the Magdalena River Basin in Clombia. *Front. Earth Sci.* **2018**, *6*, 68. [CrossRef]

3. Jia, S.; Zhu, W.; Lu, A.; Yan, T. A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sens. Environ.* **2011**, *115*, 3069–3079. [CrossRef]

4. Li, M.; Shao, Q. An improved statistical approach to merge satellite rainfall estimates and rain gauge data. *J. Hydrol.* **2010**, *385*, 51–64. [CrossRef]

5. Goodrich, D.; Faurès, J.; Woolhiser, D.; Lane, L.; Sorooshian, S. Measurement and analysis of small-scale convective storm rainfall variability. *J. Hydrol.* **1995**, *173*, 283–308. [CrossRef]

6. Wheater, H.; Isham, V.; Cox, D.; Chandler, R.; Kakou, A.; Northrop, P.; Oh, L.; Onof, C.; Rodriguez-Iturbe, I. Spatial-temporal rainfall fields: Modelling and statistical aspects. *Hydrol. Earth Syst. Sci.* **2000**, *4*, 581–601. [CrossRef]

7. Gruber, A.; Levizzani, V. *Assessment of Global Precipitation Products: A Project of the World Climate Research Programmed Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel*; WCRP-128; WMO: Geneva, Switzerland, 2008; p. 50.

8. Wilheit, T. Some comments on passive microwave measurement of rain. *Bull. Am. Meteorol. Soc.* **1986**, *67*, 1226–1232. [CrossRef]

9. Loukas, A.; Vasiliades, L. Streamflow simulation methods for ungauged and poorly gauged watersheds. *Nat. Hazard Earth Syst. Sci.* **2014**, *14*, 1641. [CrossRef]

10. Sivapalan, M.; Takeuchi, K.; Franks, S.; Gupta, V.; Karambiri, H.; Lakshmi, V.; Liang, X.; McDonnell, J.; Mendiondo, E.; O'connell, P. IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **2010**, *48*, 857–880. [CrossRef]

11. Beesley, C.; Frost, A.; Zajaczkowski, J. A comparison of the BAWAP and SILO spatially interpolated daily rainfall datasets. In Proceedings of the 18th World IMACS/MODSIM Congress, Cairns, Australia, 13–17 July 2009.

12. Hughes, D. Comparison of satellite rainfall data with observations from gauging station networks. *J. Hydrol.* **2006**, *327*, 399–410.

13. Jeffery, S. *Error Analysis for the Interpolation of Monthly Rainfall Used in the Generation of SILO Rainfall Datasets*; Technical Report; The Queensland Department of Natural Resources: Brisbane City, Australia, 2006.

14. Collischonn, B.; Collischonn, W.; Tucci, C.E.M. Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *J. Hydrol.* **2008**, *360*, 207–216. [CrossRef]

15. Bohnenstengel, S.I.; Schlünzen, K.; Beyrich, F. Representativity of in situ precipitation measurements—A case study for the LITFASS area in North-Eastern Germany. *J. Hydrol.* **2011**, *400*, 387–395. [CrossRef]

16. Dingman, S. *Physical Hydrology*; Prentice Hall: New Jersey, NJ, USA, 2002.

17. Michaelides, S.; Levizzani, V.; Anagnostou, E.; Bauer, P.; Kasparis, T.; Lane, J. Precipitation: Measurement, remote sensing, climatology and modeling. *Atmos. Res.* **2009**, *94*, 512–533. [CrossRef]

18. Nastos, P.; Kapsomenakis, J.; Philandras, K. Evaluation of the TRMM 3B43 gridded precipitation estimates over Greece. *Atmos. Res.* **2016**, *169*, 497–514. [CrossRef]

19. Adler, R.; Kidd, C.; Petty, G.; Morissey, M.; Goodman, H. Inter-comparison of global precipitation products: The third Precipitation Inter-comparison Project (PIP-3). *Bull. Am. Meteorol. Soc.* **2001**, *82*, 1377–1396.

20. Huffman, G.; Adler, R.; Arkin, P.; Chang, A.; Ferraro, R.; Gruber, A.; Janowiak, J.; McNab, A.; Rudolf, B.; Schneider, U. The global precipitation climatology project (GPCP) combined precipitation dataset. *Bull. Am. Meteorol. Soc.* **1997**, *78*, 5–20. [CrossRef]

21. Huffman, G.; Adler, R.; Morrissey, M.; Bolvin, D.; Curtis, S.; Joyce, R.; McGavock, B.; Susskind, J. Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.* **2001**, *2*, 36–50. [CrossRef]

22. Huffman, G.; Adler, R.; Bolvin, D.; Gu, G. Improving the global precipitation record: GPCP version 2.1. *Geophys. Res. Lett.* **2009**, *36*. [CrossRef]

23. Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Hirose, M.; Takayabu, Y.; Ushio, T.; Nakagawa, K. Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Trans. Geosci. Remote Sens. Environ.* **2007**, *45*, 2259–2275. [CrossRef]

24. Beck, H.; Van Dijk, A.; Levizzani, V.; Schellekens, J.; Gonzalez Miralles, D.; Martens, B.; De Roo, A. MSWEP: 3-hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 589–615. [CrossRef]

25. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 1–21. [CrossRef]

26. Hsu, K.-L.; Gao, X.; Sorooshian, S.; Gupta, H. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.* **1997**, *36*, 1176–1190. [CrossRef]

27. Kummerow, C.; Barnes, W.; Kozu, T.; Shiue, J.; Simpson, J. The tropical rainfall measuring mission (TRMM) sensor package. *J. Atmos. Ocean. Technol.* **1998**, *15*, 809–817. [CrossRef]

28. Kummerow, C.; Simpson, J.; Thiele, O.; Barnes, W.; Chang, A.; Stocker, E.; Adler, R.; Hou, A.; Kakar, R.; Wentz, F. The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteorol.* **2000**, *39*, 1965–1982. [CrossRef]

29. Huffman, G.; Bolvin, D.; Nelkin, E.; Wolff, D.; Adler, R.; Gu, G.; Hong, Y.; Bowman, K.; Stocker, E. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]

30. Ma, Z.; Shi, Z.; Zhou, Y.; Xu, J.; Yu, W.; Yang, Y. A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed. *Remote Sens. Environ.* **2017**, *200*, 378–395. [CrossRef]

31. Mahmud, M.; Mohd Yusof, A.; Mohd Reba, M.; Hashim, M. Mapping the daily rainfall over an ungauged tropical micro-watershed: A downscaling algorithm using GPM data. *Water* **2020**, *12*, 1661.

32. Hou, A.; Kakar, R.; Neeck, S.; Azarbarzin, A.; Kummerow, C.; Kojima, M.; Oki, R.; Nakamura, K.; Iguchi, T. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 701–722. [CrossRef]

33. Huffman, G.; Stocker, E.; Bolvin, D.; Nelkin, E.; Tan, J. GPM IMERG Final Precipitation L3 1 day 0.1 degree × 0.1 degree V06. In *Goddard Earth Sciences Data and Information Services Center (GES DISC)*; Andrey, G., Savtchenko, M.D., Eds.; NASA: Washington, DC, USA, 2019.

34. Duan, Z.; Bastiaanssen, W. First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure. *Remote Sens. Environ.* **2013**, *131*, 1–13. [CrossRef]

35. Hunink, J.; Immerzeel, W.; Droogers, P. A High-resolution Precipitation 2-step mapping Procedure (HiP2P): Development and application to a tropical mountainous area. *Remote Sens. Environ.* **2014**, *140*, 179–188. [CrossRef]

36. Zhang, Q.; Shi, P.; Singh, V.; Fan, K.; Huang, J. Spatial downscaling of TRMM-based precipitation data using vegetative response in Xinjiang, China. *Int. J. Climatol.* **2017**, *37*, 3895–3909. [CrossRef]

37. Ulloa, J.; Ballari, D.; Campozano, L.; Samaniego, E. Two-step downscaling of TRMM 3B43 V7 precipitation in contrasting climatic regions with sparse monitoring: The case of Ecuador in Tropical South America. *Remote Sens.* **2017**, *9*, 758. [CrossRef]

38. Shi, Y.; Song, L.; Xia, Z.; Lin, Y.; Myneni, R.; Choi, S.; Wang, L.; Ni, X.; Lao, C.; Yang, F. Mapping annual precipitation across Mainland China in the period 2001–2010 from TRMM 3B43 product using spatial downscaling approach. *Remote Sens.* **2015**, *7*, 5849–5878. [CrossRef]

39. Immerzeel, W.; Rutten, M.; Droogers, P. Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sens. Environ.* **2009**, *113*, 362–370. [CrossRef]

40. Chen, F.; Liu, Y.; Liu, Q.; Li, X. Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity. *Int. J. Remote Sens.* **2014**, *35*, 3074–3093. [CrossRef]

41. Alexakis, D.; Tsanis, I. Comparison of multiple linear regression and artificial neural network models for downscaling TRMM precipitation products using MODIS data. *Environ. Earth Sci.* **2016**, *75*, 1077. [CrossRef]

42. Ceccherini, G.; Ameztoy, I.; Hernández, C.; Moreno, C. High-resolution precipitation datasets in South America and West Africa based on satellite-derived rainfall, enhanced vegetation index and digital elevation model. *Remote Sens.* **2015**, *7*, 6454–6488. [CrossRef]

43. Goovaerts, P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **2000**, *228*, 113–129. [CrossRef]

44. Zhang, Y.; Li, Y.; Ji, X.; Luo, X.; Li, X. Fine-resolution precipitation mapping in a mountainous watershed: Geostatistical downscaling of TRMM products based on environmental variables. *Remote Sens.* **2018**, *10*, 119. [CrossRef]

45. Fang, J.; Du, J.; Xu, W.; Shi, P.; Li, M.; Ming, X. Spatial downscaling of TRMM precipitation data based on the orographical effect and meteorological conditions in a mountainous area. *Adv. Water Resour.* **2013**, *61*, 42–50. [CrossRef]

46. Park, N. Spatial downscaling of TRMM precipitation using geostatistics and fine scale environmental variables. *Adv. Meteorol.* **2013**, *2013*, 237126. [CrossRef]

47. Chen, C.; Zhao, S.; Duan, Z.; Qin, Z. An improved spatial downscaling procedure for TRMM 3B43 precipitation product using geographically weighted regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4592–4604. [CrossRef]

48. Xu, G.; Xu, X.; Liu, M.; Sun, A.; Wang, K. Spatial downscaling of TRMM precipitation product using a combined multifractal and regression approach: Demonstration for South China. *Water* **2015**, *7*, 3083–3102. [CrossRef]

49. Ezzine, H.; Bouziane, A.; Ouazar, D.; Hasnaoui, M. Downscaling of open coarse precipitation data through spatial and statistical analysis, integrating NDVI, NDWI, ELEVATION, and distance from sea. *Adv. Meteorol.* **2017**, *2017*, 8124962. [CrossRef]

50. Xu, S.; Wu, C.; Wang, L.; Gonsamo, A.; Shen, Y.; Niu, Z. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. *Remote Sens. Environ.* **2015**, *162*, 119–140. [CrossRef]

51. Mitas, L.; Mitasova, H. General variational approach to the interpolation problem. *Comput. Math. Appl.* **1988**, *16*, 983–992. [CrossRef]

52. Li, S.; Yang, S.; Liu, X. Spatiotemporal variability of extreme precipitation in north and south of the Qinling-Huaihe region and influencing factors during 1960–2013. *Prog. Geogr.* **2015**, *34*, 354–363.

53. NASA; METI; AIST; Japan Space System. ASTER Global Digital Elevation Model V003. NASA EOSDIS Land Processes DAAC. 2019. Available online: https://scholar.smu.edu/datasciencereview/vol1/iss3/9 (accessed on 20 June 2020).

54. Ullah, S.; Farooq, M.; Sarwar, T.; Tareen, M.J.; Wahid, M.A. Flood modeling and simulations using hydrodynamic model and ASTER DEM—A case study of Kalpani River. *Arab. J. Geosci.* **2016**, *9*, 439. [CrossRef]

55. Huffman, G.; Adler, R.; Bolvin, D.; Nelkin, E. The TRMM Multi-satellite Precipitation Analysis (TMPA), in Chapter 1. In *Satellite Rainfall Applications for Surface Hydrology*; Springer: Dordrecht, The Netherlands, 2010.

56. Miller, R.; Siegmund, D. Maximally selected Chi-square statistics. *Biometrics* **1982**, *38*, 1101–1106. [CrossRef]

57. Ahn, C.W.; Ramakrishna, R.S. Elitism-based compact genetic algorithms. *IEEE Trans. Evol. Comput.* **2003**, *4*, 367–385.

58. Zuo, Z.; Yan, L.; Ullah, S.; Sun, S.; Zhang, R.; Zhao, H. Empirical distribution based framework for improving multi-parent crossover algorithms. *Soft Comput.* **2020**, in press.

59. Verlinde, J. *TRMM Rainfall Data Downscaling in the Pangani Basin in Tanzania*; Delft University of Technology: Delft, The Netherlands, 2011.

60. Franke, R. Smooth interpolation of scattered data by local thin plate splines. *Comput. Math. Appl.* **1982**, *8*, 237–281. [CrossRef]

61. Eiben, A.E.; Back, T. Empirical investigation of multiparent recombination operators in evolution strategies. *Evol. Comput.* **1997**, *5*, 347–365. [CrossRef] [PubMed]

62. Herrera, F.; Lozano, M.; Verdegay, J.L. Tackling real-coded genetic algorithms: Operators and tools for behavioral analysis. *Artif. Intell. Rev.* **1998**, *12*, 265–319. [CrossRef]

63. Goldberg, D.E. Real-coded genetic algorithms, virtual alphabets, and blocking. *Complex Syst.* **1991**, *5*, 139–167.

64. Agam, N.; Kustas, W.P.; Anderson, M.C.; Li, F.; Neale, C.M.U. A vegetation index based 572 technique for spatial sharpening of thermal imagery. *Remote Sens. Environ.* **2007**, *107*, 545–558. [CrossRef]

MDPI

Article

Capacity of Satellite-Based and Reanalysis Precipitation Products in Detecting Long-Term Trends across Mainland China

Shanlei Sun [1],*, Wanrong Shi [1], Shujia Zhou [2], Rongfan Chai [1], Haishan Chen [1], Guojie Wang [3], Yang Zhou [1] and Huayu Shen [4]

[1] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory of Meteorological Disaster, Ministry of Education/International Joint Research Laboratory on Climate and Environment Change, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China; 20181201117@nuist.edu.cn (W.S.); rfchai@nuist.edu.cn (R.C.); haishan@nuist.edu.cn (H.C.); yangzhou@nuist.edu.cn (Y.Z.)

[2] Jintan Meteorological Bureau, Changzhou 213200, China; 20191201122@nuist.edu.cn

[3] School of Geographical Sciences, NUIST, Nanjing 210044, China; gwang@nuist.edu.cn

[4] Ningbo Meteorological Bureau, Ningbo 315012, China; qxj@nbyz.gov.cn

* Correspondence: sun.s@nuist.edu.cn; Tel.: +86-025-5869-5622

Abstract: Despite numerous assessments of satellite-based and reanalysis precipitation across the globe, few studies have been conducted based on the precipitation linear trend (LT), particularly during daytime and nighttime, when there are different precipitation mechanisms. Herein, we first examine LTs for the whole day (LT_{wd}), daytime (LT_d), and nighttime (LT_n) over mainland China (MC) in 2003–2017, with sub-daily observations from a dense rain gauge network. For MC and ten Water Resources Regions (WRRs), annual and seasonal LT_{wd}, LT_d, and LT_n were generally positive but with evident regional differences. Subsequently, annual and seasonal LTs derived from six satellite-based and six reanalysis popular precipitation products were evaluated using metrics of correlation coefficient (CC), bias, root-mean-square-error (RMSE), and sign accuracy. Finally, metric-based optimal products (OPs) were identified for MC and each WRR. Values of each metric for annual and seasonal LT_{wd}, LT_d, or LT_n differ among products; meanwhile, for any single product, performance varied by season and time of day. Correspondingly, the metric-based OPs varied among regions and seasons, and between daytime and nighttime, but were mainly characterized by OPs of Tropical Rainfall Measuring Mission (TRMM) 3B42, ECMWF Reanalysis (ERA)-Interim, and Modern Era Reanalysis for Research and Applications (MERRA)-2. In particular, the CC-based (RMSE-based) OPs in southern and northern WRRs were generally TRMM3B42 and MERRA-2, respectively. These findings imply that to investigate precipitation change and obtain robust related conclusions using precipitation products, comprehensive evaluations are necessary, due to variation in performance within one year, one day and among regions for different products. Additionally, our study facilitates a valuable reference for product users seeking reliable precipitation estimates to examine precipitation change across MC, and an insight (i.e., capacity in detecting LTs, including daytime and nighttime) for developers improving algorithms.

Keywords: precipitation; reanalysis; satellite; linear trends; mainland China

---

1. Introduction

Precipitation is a critical hydrometeorogical variable that plays a key role in energy and water cycles, and thus impacts the weather, climate, hydrology, ecosystem, and Earth system [1–3]. Precipitation is closely bound to life on Earth, due to it being the major source of freshwater [4]. As a result,

its measurement is the focus for various disciplines (e.g., atmospheric sciences, ecology, hydrology, agriculture, and economy), in spite of their differences [2,5–7]. For instance, continuous and long-term precipitation observations are necessary for scientific research, but also for informing policy-makers of suitable measures to mitigate the adverse impacts of climate change, especially for droughts and floods [8–12]. Despite its exceptional importance, there still exist great challenges related to obtaining reliable precipitation observations with a long enough time span and large enough space coverage [13–15].

It is well-known that the most direct pathway to obtain precipitation data is through in situ measurements with different gauges (e.g., tipping-bucket rain gauges; [16]). Gauge observations have been extensively utilized by different sectors (e.g., agriculture, industry, and forestry), and have greatly promoted precipitation-related scientific disciplines, e.g., climate studies [10,17,18]. Despite that, we should note that gauges are related to high variability of the rain-bearing systems at different spatio-temporal scales and have an uneven spatial distribution [19,20]. These factors limit the representativeness of gauge precipitation observations to a large extent, and introduce uncertainty into gauge data-based conclusions. With the development of radar technology over past decades, technologically-sophisticated precipitation algorithms based on radar radiance signals have been developed and various corresponding precipitation estimates have been proposed [21–23]. Radar-based precipitation products undoubtedly have the potential to solve or at least reduce the limitations of gauge measurements; for example, they have a more extensive coverage and higher spatio-temporal resolution, which are critical for the analysis of hydrometeorological processes, especially extremes such as flash floods and droughts [2,24]. However, considering the high installation and maintenance costs of radars, the shorter time span of radar observations, and topography-induced radar blockage, radar-based precipitation products are limited and unavailable for some regions, e.g., the nearly negligible radar estimates overseas [14,22,25].

Recently, satellite technology and numerical models/reanalysis systems have developed significantly; therefore, satellite-based precipitation retrievals and precipitation estimates from the models/reanalysis systems have become increasingly attractive and available [3,5,26–31]. Satellite-based precipitation data are derived using various statistical and/or physics-based retrieval algorithms with the radiance information from the satellite-carried sensors, including VIS/IR sensors on geostationary (GEO) and low Earth orbit (LEO) satellites, and passive (PMW) and active MW sensors on LEO satellites. Satellite-based products generally include three types, i.e., VIS/IR- and MW-based estimates, and data through combining the VIS/IR and MW information. Examples include the Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA) from the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC; [32]), the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; [33,34]), the PERSIANN-Cloud Classification System (CCS; [13,35]), the PERSIANN-Climate Data Record (CDR; [25]), the Global Satellite Mapping of Precipitation (GSMaP) Microwave-Infrared Combined Reanalysis (RNL; [32,36]) of the Earth Observation Research Center (EORC) of the Japan Aerospace Exploration Agency (JAXA), the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) morphing technique (CMORPH; [37]), the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (IMERG; [38]), the Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS; [39]), and a new global Multi-Source Weighted-Ensemble Precipitation (MSEWP) rainfall dataset [40]. Notably, satellite-based precipitation products can only be dated back to the beginning of the satellite era. As an important alternative, reanalysis precipitation data, which are simulated using new atmospheric models combined with advanced data assimilation systems [41–45], provide a means to fulfill the specific requirements (i.e., centennial and even longer precipitation records) of the studies. Additionally, reanalysis systems provide the nearly realistic atmosphere circulation fields, which makes it possible to do many things, including understanding precipitation changes from the perspective of atmospheric dynamical mechanisms [46–48]. The most frequently used reanalysis precipitation datasets include the National Centers for Environmental Prediction reanalysis (NCEP1 and NCEP2; [42,43]), NCEP Climate Forecast System Reanalysis (CFSR; [49]),

ECMWF Reanalysis-5 (ERA-5; [46]) and ERA-Interim [41], the Japanese 55-year Reanalysis Project (JRA-55; [44]), and the National Aeronautics and Space Administration (NASA) Modern Era Reanalysis for Research and Applications (e.g., MERRA-2; [45]).

Before using these products, it is of paramount importance to determine the reliability of the precipitation products using dependable reference datasets, because the inherent uncertainties within these products would likely affect final results, adversely impacting confidence levels [42,50–56]. In terms of a study's specific needs and goals, the satellite-based and reanalysis precipitation datasets have been widely evaluated at different spatio-temporal scales with a series of validation metrics (e.g., [6,50,56–68]). For instance, Sun et al. [6] selected several continuous and categorical validation statistics combined with bias and error decomposition techniques to assess the performance of the PERSIANN-Climate Data Record (CDR) precipitation product in the Huai River Basin, China, and pointed out that the daily, monthly and annual performance of this product varied in accordance with obvious intra-annual cycles. Huang et al. [62] systematically assessed five satellite-based precipitation products (CMORPH, PERSIANN and TRMM3B41RT, TRMM3B42RT, and TRMM3B42) with observations at 2400 weather sites across China, and found that estimates generally captured the overall spatial-temporal variation of precipitation, especially for warm seasons and humid regions. Beck et al. [40] compared 22 gridded daily precipitation datasets across the globe during 2000—2016 with daily observations at 76086 gauges and hydrological modeling, and highlighted that there existed large differences in the accuracy of precipitation estimates and more attention should be paid for precipitation dataset selection in both research and operational applications. de Leeuw et al. [65] used the daily precipitation observations from England and Wales to evaluate the ERA-Interim products, and found that this dataset underestimated the observations on a daily scale, while it could capture the statistics of extreme precipitation events. Lorenz and Kunstamann [67] analyzed the hydrological cycle with three state-of-the-art reanalyses (ERA-Interim, MERRA-2, and CFSR), and demonstrated that large differences existed between the reanalyses and the observations.

The previous evaluations have provided valuable information for the theoretical understanding and improvement of satellite-retrieved algorithms and reanalysis systems. Nonetheless, most were conducted using daily, monthly and annual reference precipitation data; thus, the information about the capacity of the satellite-based and reanalysis precipitation is scarce on a sub-daily scale, especially for China. In fact, there are evident differences in the mechanisms of precipitation within one day, which are closely related to thermodynamic and dynamic processes of water and energy fluxes [3,69–74]. For example, results from Yu et al. [74] indicated that long-duration stratiform precipitation frequently occurred in the early morning during the warm season over central-eastern China, while the late afternoon experienced a higher frequency of short-duration convective precipitation. Therefore, evaluating the multi-source precipitation products with sub-daily observations (daytime and nighttime datasets at least) could provide more detailed information, e.g., flexibility for a precipitation product on sub-daily scale. This is very useful to further improve satellite-based algorithms and models/reanalysis systems from the perspective of sub-daily precipitation mechanisms, and even correct the precipitation products using the sub-daily rather than daily measurements. Additionally, sub-daily precipitation changes have become a hot topic in current research, and numerous studies have been conducted (e.g., [71,75–82]). Cheng et al. [71] pointed out that on annual and seasonal scales (except during spring), the majority of meteorological station records (1961–2006) of Southwest China displayed downward trends for total, daytime, and nighttime precipitation. Lin et al. [80] analyzed characteristics of summer precipitation diurnal variations during 2001–2014 in the Hubei Province of China, and suggested that the diurnal variations existed obvious regional differences. Based on observational day and night precipitation during 1961–2005 across Xinjiang, China, Han et al. [81] concluded that the annual increasing trends of precipitation in the daytime and nighttime respectively accounted for 49% and 51% of the total increasing trend in annual precipitation. Liu et al. [82] found that with the CMORPH dataset during 2008–2014, both the daytime and nighttime precipitation were detected to increase in summer over the Qilian Mountains, China. Lenderink et al. [77] reported that hourly precipitation

extremes have substantially increased in the last century over De Bilt, Netherlands, and Hong Kong, China. Thus, an issue arises—can the existing precipitation products capture the linear trends on a sub-daily scale based on different validation metrics? This question has been paid little attention (e.g., [83]), despite the basis to examine precipitation trends with these datasets. Thus, assessments regarding precipitation trends can provide fundamental information to select the reliable products for exploring precipitation changes, particularly for regions with limited or even no observations (e.g., West China in Figure 1).

Considering the gaps in the previous works of precipitation evaluations, we used China as an example to examine the multi-source precipitation products' capacity to detect precipitation linear trends during daytime and nighttime. Thus, the main objectives of this work were to (1) investigate the spatial distribution of precipitation changes using daily, daytime, and nighttime records from 2393 weather sites across China; (2) to quantify the performance of selected products (i.e., six satellite-based and six reanalysis datasets) in detecting precipitation trends on a sub-daily scale with different validation metrics (correlation coefficient, bias, root mean square error, and sign accuracy) through a comparison with gauge observations; and (3) to identify the metric-based optimal products at a sub-daily scale.

2. Data and Methodology

2.1. Data

2.1.1. Observed Precipitation

To evaluate the capabilities of various products in capturing precipitation changes, sub-daily (i.e., daytime, P_d, 0000–1200 UTC; and nighttime, P_n, 1200–2400 UTC) accumulated precipitation data observed from 2003 to 2017 at 2481 weather sites across China (Figure 1), including basic, benchmark, and general meteorological stations, were collected from the China Meteorological Administration (CMA). Although both datasets had undergone a series of quality control measures and homogenization, e.g., outlier identification, internal consistency checks, and spatio-temporal consistency checks [84], there were still missing values within the records. Therefore, to maximize the observational information, we processed the datasets following the procedures described below. First, the number of daytime and nighttime values was computed for each year at each site. If the days with missing values for daytime or nighttime observations exceed 50 at a site, the site was removed. Secondly, for the remaining sites, the bilinear interpolation method was employed to fill the missing values with the observations at the two closest sites. There were 2393 sites remaining after this process (Figure 1), and the accumulative precipitation for a whole day (abbreviated as P_{wd}) was then obtained as the sum of P_d and P_n.
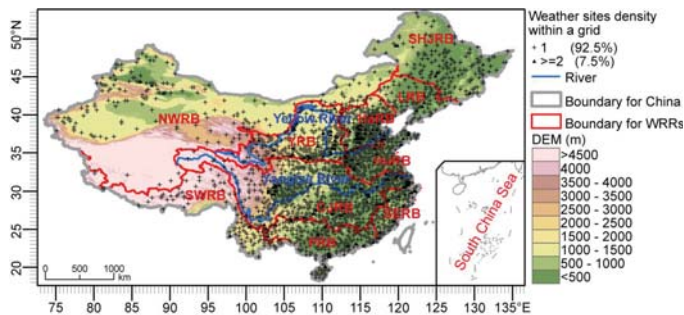
Figure 1. Geographic distribution of the selected 2281 grids (0.25° × 0.25°), at least corresponding to a weather site. The digital elevation model (DEM) with a spatial resolution of 90 m is available at http://srtm.csi.cgiar.org/. [85] Crosses and triangles correspond to 1 and more than 2 sites within a given grid, respectively, followed by the percentage of grid shown in the bracket.

China is located in a typical monsoon region (i.e., the East Asian monsoon region), with evident spatio-temporal variability of precipitation and the related mechanisms [86]. Mainland China (MC) is divided into ten Water Resources Regions (WRRs, Figure 1), which is beneficial for examining regional differences in the performance of each product to detect precipitation trends. We conducted annual and seasonal evaluations of P_{wd}, P_d, and P_n trends during 2003–2017 on national and regional (i.e., MC and WRR, respectively) scales. Here, spring, summer, autumn, and winter were specified as March–May, June–August, September–November, and December–February, respectively.

2.1.2. Satellite-Based and Reanalysis Precipitation Datasets

In this study, considering the precipitation datasets availability and time span (study period of 2004–2017), we collected twelve sets of gridded precipitation data, including six satellite-based and six reanalysis products, for evaluation. Detailed information of these datasets is shown in Table 1. Of the selected satellite-based precipitation products, both the TRMM (i.e., TRMM3B42RT and TRMM3B42 adjusted with gauge observations), and the GSMaP (i.e., GSMaP-RNL and GSMaP-RNLG adjusted with gauge observations) precipitation datasets are produced through merging VIS/IR and MW information but are based on different algorithms [32,36]. In contrast, PERSIANN and PERSIANN-CCS belong to the VIS/IR family of satellite-based precipitation products [13,33–35,87]. The main differences in the two PERSIANN products are that the PERSIANN-CCS system enables the categorization of cloud-patch features based on cloud height, areal extent, and variability of texture estimated from satellite imagery, which is optimized for observing extreme precipitation, particularly at a very high spatial resolution. The six reanalysis precipitation products include JRA-55, ERA-Interim, ERA-5, NCEP1, NCEP2, and MERRA-2. These reanalysis products are produced based on different forecasting systems by assimilating many of the basic surface and upper-atmospheric fields from multiple sources, e.g., the surface humidity, radiosonde-based specific humidity, wind fields, and satellite-derived radiance. Among them, different data assimilation techniques are employed. For example, the ERA-Interim, ERA-5, and JRA-55 adopt four-dimensional variational (4D-VAR) data assimilation systems, whereas the MERRA-2, NCEP1, and NECEP2 utilize 3D-VAR assimilation systems. For more details about these datasets, the reader can refer to the product-specific user guide and the related literature.

Table 1. Summary of the selected satellite and reanalysis rainfall products.

Products	Spatial Resolution and Space Span	Temporal Resolution and Time Span	Bias Correction	Assimilation System	References
TRMM-3B42RT (V7)	0.25° × 0.25°, 50° S–50° N	2000 to present, 3-hourly	No	/	[32]
TRMM-3B42 (V7)	0.25° × 0.25°, 50° S–50° N	2000 to 2017, 3-hourly	Corrected with GPCP, and CAMS	/	[32]
PERSIANN	0.25° × 0.25°, 60°S–60°N	2000 to present, 3-hourly	No	/	[34]
PERSIANN-CCS	0.04° × 0.04°, 60° S–60° N	2003 to present, 3-hourly	No	/	[35]
GSMaP-RNL (V6)	0.1° × 0.1°, 60° S–60° N	2000 to present, hourly	No	/	[36]
GSMaP-RNLG (V6)	0.1° × 0.1°, 60° S–60° N	2000 to present, hourly	Corrected with CPCU	/	[36]
JRA-55	1.25° × 1.25°, Global	1958 to present, 3-hourly	No	4D-VAR	[44]
ERA-Interim	0.75° × 0.75°, Global	1979 to present, 3-hourly	No	4D-VAR	[49]
ERA-5	0.25° × 0.25°, Global	1979 to present, 3-hourly	No	4D-VAR	[46]
NCEP1	1.875° × 1.875°, Global	1948 to present, 6-hourly	No	3D-VAR	[42]
NCEP2	1.875° × 1.875°, Global	1979to present, 6-hourly	No	3D-VAR	[43]
MERRA-2	0.5° × 0.667°, Global	1980 to present, hourly	Corrected with CPCU or CMAP/GPCPv2.1	3D-VAR	[45]

Note: CPCU: the NOAA Climate Prediction Center (CPC) Unified Gauge-Based Analysis of Global Daily Precipitation (CPCU) product. CMAP/GPCPv2.1: the CPC Merged Analysis of Precipitation (CMAP)/Global Precipitation Climatology Project product, version 2.2. GPCC: Global Precipitation Climatology Centre. CAMS: Climate Assessment and Monitoring Systems.

As shown in Table 1, datasets had different temporal and spatial resolutions, so it is necessary to process them before evaluation. First, the satellite-based and reanalysis P_d and P_n were summed from the 1-hourly, 3-hourly, or 6-hourly accumulated precipitation at product-specified grids. Then, based on the bilinear interpolation method, the P_d and P_n for all products (except for TRMM-3B42RT, TRMM-3B42, PERSIANN, and ERA-5) were resampled to the spatial resolution of $0.25°$. This was mainly because most products correspond to a spatial resolution of $0.25°$ or higher, so the resampling-induced uncertainties could be reduced to some extent. For P_{wd}, its values were obtained using the sum of P_d and P_n from the resampled maps. The grids with at least one site were extracted to conduct performance evaluations. If any grid included more than one site, the average precipitation value at these sites was calculated to represent the final reference value of that grid.

2.2. Methodolody

The precipitation trends were calculated using

$$y = at + b \tag{1}$$

where y is annual or seasonal accumulative precipitation; t refers to time; a represents the slope coefficient, namely, linear trend; and b is the constant. Pearson's correlation and the two-tailed Student's t test (i.e., $p < 0.05$) were applied to check for statistically significant relationships.

Satellite-based and reanalysis precipitation trends were quantitatively assessed with the metrics of bias (B), which measured the trend differences between the products and the gauge observations; root mean square error (RMSE), which represented the overall accuracy of the trends derived from the products; the correlation coefficient (CC), which quantified the spatial consistency of the trends derived from the products; and accuracy of sign (AS), which examined the degree of agreement between the positive or negative sign of precipitation trends from the products and the observed data. These metrics were calculated using the following equations:

$$B = \overline{a_P} - \overline{a_O} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (a_{O,i} - a_{P,i})^2} \tag{3}$$

$$CC = \frac{\sum_{i=1}^{N} (a_{O,i} - \overline{a_O})(a_{P,i} - \overline{a_P})}{\sqrt{\sum_{i=1}^{N} (a_{O,i} - \overline{a_O})^2 \sum_{i=1}^{N} (a_{P,i} - \overline{a_P})^2}} \tag{4}$$

$$AS = \frac{nP}{nG} \tag{5}$$

where $a_{P,i}$ and $a_{O,i}$ represent the linear trends from a certain precipitation product and the gauge observation at the ith grid, respectively; N is the number of the used grids for evaluation across MC or each WRR; $\overline{a_P}$ and $\overline{a_O}$ represent the products and the observed trends averaged at the grids within MC or a certain WRR, respectively; and nP is the number of the grids, where the examined products shows the same sign of precipitation (e.g., P_{wd}, P_d or P_n) changes as the observed within a given region, but nG indicates the total number of grids in the region.

Considering the co-variation of P_d and P_n, we defined a joint AS (JAS), which represented the capacity of a given product to rightly detect the signs of both P_d and P_n changes relative to the observed data. JAS can be calculated by

$$JAS = \frac{nP_{co}}{nG} \tag{6}$$

where nP_{co} is the number of the grids in which the signs of both P_d and P_n changes derived from the products are the same as those observed in a given region.

3. Results

3.1. Gauge Precipitation Changes across MC

Figure 2(a1) depicts observed annual LTs for MC and ten WRRs during 2003–2017. For MC, annual LT_{wd} and LT_d were 8.42 mm/yr ($p < 0.05$) and 4.96 mm/yr ($p < 0.05$), respectively, followed by an insignificant LT_n of 3.46 mm/yr. Comparing LT_{wd}, LT_d, or LT_n (i.e., signs and magnitudes) among WRRs, there were evident regional differences, while significant ($p < 0.05$) and larger increases (>13 mm/yr, >8 mm/yr and >7 mm/yr for LT_{wd}, LT_d and LT_n, respectively) were found in YZRB, SERB, and PRB, followed by the largest reductions (–13.95 mm/yr for LT_{wd}, −4.68 mm/yr for LT_d and −9.27 mm/yr for LT_n) in HuRB. In spring (Figure 2(b1)), LTs for WRRs and MC were between −4 mm/yr and 4 mm/yr, with the exceptions of SERB and PRB, which showed LT_{wd} > 8 mm/yr, and LT_d and LT_n > 4 mm/yr. During summer, MC LT_{wd} and LT_d (LT_n) were positive (negative) with a rate <2 mm/yr (Figure 2(c1)). Among ten WRRs, most exhibited smaller LT_{wd} (LT_d and LT_n) in summer, generally corresponding to between −3 mm/yr and 4 mm/yr (between −1.50 mm/yr and 2.50 mm/yr); however, significant ($p < 0.05$) decreasing and increasing LTs were detected over HuRB and YZRB (excluding LT_n) and SERB, and the LT_{wd}, LT_d, and LT_n were >6 mm/yr and > 4 mm/yr, respectively. As shown in Figure 2(d1), MC LT_{wd} and LT_d (LT_n) were 5.60 mm/yr ($p < 0.05$) and 2.80 mm/yr ($p < 0.05$), respectively. Except for two WRRs (i.e., HaRB and YRB), autumn LT_{wd}, LT_d, and LT_n were consistently positive from 2003 to 2017. However, magnitudes of autumn LTs differed among these WRRs, for which significant ($p < 0.05$) and larger increases (>8 mm/yr for LT_{wd} and 3.20 mm/yr for LT_d and LT_n) occurred in YZRB, SERB, and PRB. Regarding winter precipitation (Figure 2(e1)), SERB and PRB exhibited the highest LT_{wd} (>2.80 mm/yr) and LT_d and LT_n > 0.90 mm/yr, followed by the remaining WRRs and MC with an LT_{wd} < 1.50 mm/yr (LT_d and LT_n < 0.70 mm/yr). Additionally, comparing signs and magnitudes of LT_d and LT_n (Figure 2(a1–e1)), 10 and 15 of 55 cases (i.e., 11 (MC + 10 WRRs) × 5 (annual + seasonal scales)) showed opposite signs and larger differences, with ratios between LT_d and LT_n > 2.00 and < 0.50, respectively. These findings imply that LT_d and LT_n values were not consistent, possibly due to the different precipitating mechanisms during daytime and nighttime, and thus further confirms the necessity to evaluate various precipitation products at a sub-daily scale.

As shown in Table 2 and Figure 2(a2), 33% of grids had decreasing annual LT_{wd} across MC, generally in east LRB, HuRB, the YRB–YZRB border, and most of SWRB and NWRB. Moreover, 3% of grids in north-central HuRB showed significant ($p < 0.05$) negative annual LT_{wd} with a rate of −12 mm/yr. In contrast, 11% of grids had significantly ($p < 0.05$) increasing LT_{wd}, mainly situated in east SHRB, central YRB, northeast YZRB, north SERB, and middle PRB, for which LT_{wd} over the three latter regions exceeded 20 mm/yr. For both annual LT_d and LT_n (Table 2), negative values were found in > 30% of grids, followed by < 4% of grids with significant ($p < 0.05$) values. Moreover, in spite of smaller magnitudes of difference compared to annual LT_{wd}, similar spatial distributions for LT_d and LT_n were detected (Figure 2(a3,a4)). Figure 2(b2–b4,c2–c4,d2–d4,e2–e4) illustrate the spatial distribution of seasonal LT_{wd}, LT_d, and LT_n during 2003–2017. In broad terms, LT_{wd}, LT_d, or LT_n spatially differ during seasons, while in a given season, a generally similar spatial pattern is observed among LT_{wd}, LT_d, and LT_n, including for locations with significant ($p < 0.05$) LTs. For example, spring LT_{wd}, LT_d, and LT_n were negative at 30% of grids, primarily in NWR, SWRB, west YZRB, north HuRB, and HaRB (Figure 2(b2–b4) and Table 2); moreover, 2% of grids with significant ($p < 0.05$) changes were sporadically distributed, and larger reductions (–6 mm/yr for LT_{wd}, but –2 mm/yr for LT_d and LT_n) were in south SWRB. At the remaining grids, 60% of grids with larger increases for LT_{wd} (12 mm/yr), LT_d and LT_n (4 mm/yr) in spring for were mainly located in east YZRB, PRB, and SERB, and 5% of grids with significant ($p < 0.05$) changes were generally in SHRB–LRB, YRB–YZRB borderlands, and east PRB. As shown in Figure 2(c2–c4) and Table 2, 44% of grids with a negative summer LT_{wd}, LT_d, and LT_n were generally situated in central SHRB, LRB, HuRB, YRB–YZRB borderlands, west YZRB, PRB, and north SWRB, and the largest and significant ($p < 0.05$) reductions (−10 mm/yr) in 5% of grids were concentrated in HuRB. Of the remaining grids (>50%), the largest (10 mm/yr) and significant

($p < 0.05$) summer LTs were detected in 4% of grids mainly in northeast YZRB and north SERB. In autumn (Figure 2(d2–d4) and Table 2), LT_{wd}, LT_d, and LT_n were at least -6 mm/yr at 30% of grids in south LRB, YRB–HaRB–HuRB and YRB–YZRB borderlands, central NWRB, and central SWRB. Of the grids with increasing LTs, 14% of grids with large (10 mm/yr) and significant ($p < 0.05$) values were situated in central SHRB, central PRB, east YZRB, parts of middle YZRB (i.e., Sichuan basin), and SERB. During winter (Figure 2(e2–e4) and Table 2), there was an approximately equal balance of grids with negative and positive LT_{wd}, LT_d, or LT_n, which was generally 4mm/yr or −4mm/yr at most grids; moreover, increasing LTs were widely distributed across east coastal WRRs, south SWR, and central YRB. Furthermore, 2% of grids with significant ($p < 0.05$) increases in winter precipitation were patchily distributed across MC.
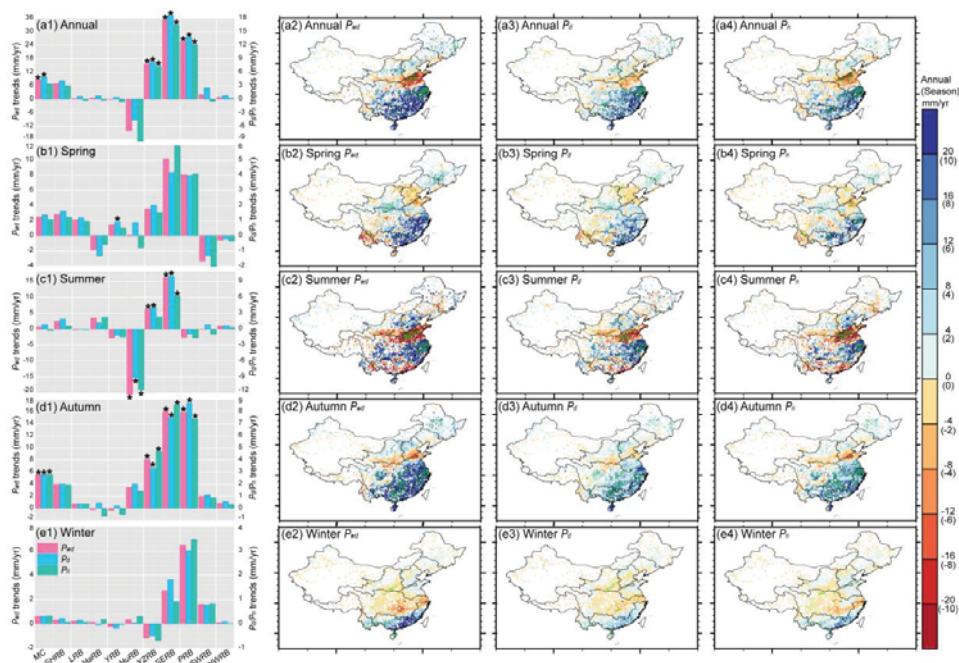
Figure 2. Linear trends (LTs) for mainland China (MC), ten Water Resources Regions (WRRs), and 2281 grids during 2003–2017. Annual and seasonal LTs averaged over MC and ten WRRs are shown in (**a1**) and (**b1–e1**), respectively, in which stars represent significant changes with $p < 0.05$. (**a2**) and (**b1–e2**) show spatial distributions of annual and seasonal P_{wd} trends across MC, respectively, with the green cross representing significant changes with $p < 0.05$. (**a1–e3**) and (**a1–e4**) are the same as (**a1–e2**), but for P_d and P_n trends, respectively.

Table 2. Percentage of grids with increasing and decreasing LTs across MC for the whole day (LT_{wd}), daytime only (LT_d), and nighttime only (LT_n).

	Annual			Spring			Summer			Autumn			Winter		
	LT_{wd}	LT_d	LT_n	LT_{wd}	LT_d	LT_n	LT_{wd}	LT_d	LT_n	LT_{wd}	LT_d	LT_n	LT_{wd}	LT_d	LT_n
Significant Increase	67	69	61	51	52	54	55	56	51	74	76	71	51	52	54
($p < 0.05$) Increase	11	11	9	6	6	5	6	6	4	17	14	17	3	2	3
Significant decrease	33	31	39	49	48	46	45	44	49	26	24	29	49	48	46
($p < 0.05$) Decrease	3	1	4	3	3	2	7	5	6	0	0	0	3	3	2

3.2. Evaluation Using Correlation Coefficient Metric

The CCs of LTs for the products and the observed values are depicted in Figure 3(a1–a5). For annual LTs, the corresponding CCs for TRMM3B42RT, TRMM3B42, PERSIANN, PERSIANN-CCS, and MERRA-2 were generally >0.40, suggesting that spatial distributions of annual LTs across MC can be derived from these products (Figure 3(a1)), especially for TRMM3B42 and MERRA-2 with CCs around 0.80. Besides, ERA-Interim, with an annual CC < 0.40, exhibited limited capacity in detecting annual LTs in space. However, annual CCs for the remaining six products were all below 0.10 and some were even negative, which indicates that these products are not able to capture the spatial distribution of LTs across MC. Comparing CCs of annual LT_d and LT_n, CC-based performance for each precipitation product differed over daytime and nighttime, especially PERSIANN-CCS and ERA-Interim, followed by TRMM3B42RT and PERSIANN. In spring (Figure 3(a2)), GSMaP-RNL, GSMaP-RNLG, JRA-55, ERA-55, NCEP1, and NCEP2 had negative CCs and therefore no ability to reflect the spatial distribution of LTs; however, the other products, with CCs > 0.40, had good performances, of which TRMM3B42 showed the best performances (CCs around 0.80) and the next was in TRMM3B42RT, PERSIANN, and MERRA-2 (CCs around 0.70). Furthermore, the spring CC-based performance of PERSIANN-CCS exhibited differences > 0.10 between daytime and nighttime. During summer (Figure 3a3), TRMM3B42 with CCs around 0.80 showed the best performance, followed by TRMM3B42RT and MERRA-2 (CCs around 0.70), ERA-Interim (CCs around 0.60), and PERSIANN and PERSIANN-CCS (CCs around 0.50). JRA-55, EAR-55, NCEP1, and NCEP2 with CCs < 0 indicated poor performance. Relative to spring, the capacity of GSMaP-RNL and GSMaP-RNLG to reproduce LTs in space increased in summer but was still limited, with CCs < 0.20. PERSIANN-CCS, GSMaP-RNL, and GSMaP-RNLG showed the greatest differences (>0.10) in summer CCs between daytime and nighttime. In autumn (Figure 3a4), the largest CCs (>0.80) were detected by TRMM3B42 and MERRA-2, while TRMM3B42RT, PERSIANN, and ERA-Interim had CCs ranging from 0.60 to 0.80. PERSIANN-CCS, JRA-55, and ERA-4 had CCs around 0.40 and could capture summer LTs spatially, while the remaining four products showed limited CC-based performance (CCs generally < 0.10). Comparisons of CCs for autumn LT_d and LT_n indicated that larger differences (>0.10) existed in PERSIANN, PERSIAN-CCS, JRA-55, and ERA-5, especially for the former three products with differences exceeding 0.20. Regarding winter CCs (Figure 3(a5)), eight of the products had values below 0.20 or 0, indicating that they had limited or no ability to capture winter LTs in space. Of the remaining products, the best product based on CC in winter was MERRA-2 (CCs around 0.90), followed by TRMM3B42 (CCs around 0.70), TRMM3B42RT (CCs around 0.50), and ERA-Interim (CCs < 0.40); no significant differences in CCs for LT_d and LT_n existed among these products.

To identify the CC-based optimal products (OPs) of LT_{wd}, LT_d, and LT_n, we compared CCs from the 12 examined products. The results are depicted in Figure 3(b1–b5). For MC, the annual, spring, summer, and autumn (excluding LT_d) CC-based OP for the three LTs was TRMM3B42, and the winter OP was MERRA-2. For annual cases (including the three LTs and ten WRRs), the CC-based OP for 17 of the 30 cases was MERRA-2, generally in northern WRRs, while 11 cases, including LTs for southern WRRs (excluding YZRB) and LT_d for LRB, HaRB, and YRB had an OP of TRMM3B42. In spring, the OP for more than ten cases was TRMM3B42, generally in southern WRRs, while 15 cases with the OP of MERRA-2 were in northern WRRs. With several exceptions (e.g., SHRB, HuRB, and NWRB) showing the summer OP of MERRA-2, TRMM3B42 was the OP in 16 cases. In winter, the OP for the overwhelming majority (27) of cases was MERRA-2, followed by three cases with ERA-Interim in SERB. Notably, some cases had CCs below 0.40 for the identified OPs, e.g., for LT_{wd}, LT_d, and LT_n in LRB and NWRB; this indicates that using the so-called CC-based OPs to represent spatial distribution of precipitation trends needs more caution in certain regions.
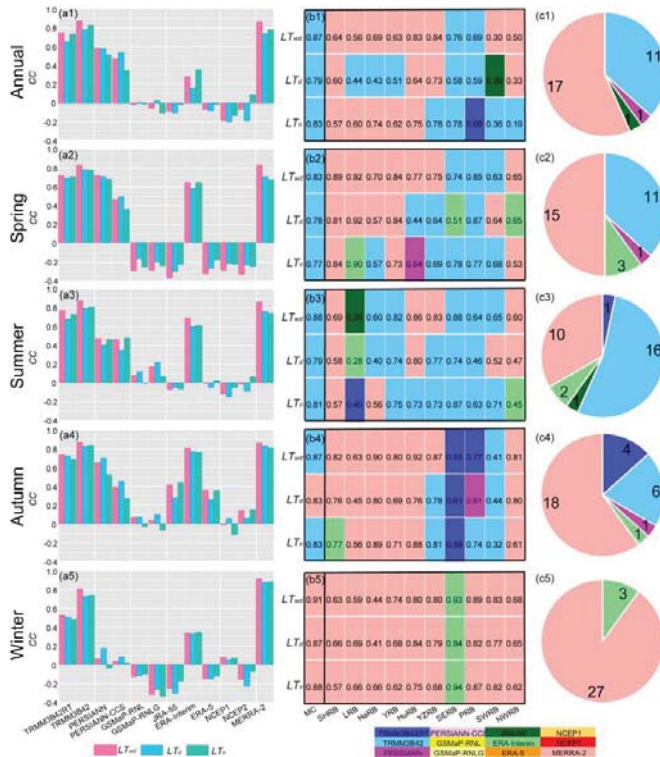
Figure 3. Correlation coefficients (CCs) for LTs from the selected 12 precipitation products (**a1–a5**), CC-based optimal products (OPs) for MC and ten WRRs (**b1–b5**), and number of cases corresponding to OPs for an annual or seasonal scale in ten WRRs (**c1–c5**). In figures (**b1–b5**), the number of each box represents the CC of the identified OP, which has been labelled with different colors. The number of figures (**c1–c5**) indicates the amount of a certain OP.

3.3. Evaluation Using Bias Metric

Figure 4(a1–a5,b1–b5) depict the percentage of grids with negative and positive Bs of annual and seasonal LTs across MC, respectively. For simplicity, we focused on analyses regarding negative Bs in this paragraph. More than 50% of grids had negative annual Bs for PERSIANN, PERSIANN-CCS, GSMaP-RNL, and GSMaP-RNLG products (Figure 4(a1)). In particular, PERSIANN LT_{wd} and LT_d had negative Bs in >65% of grids. With several exceptions (i.e., TRMM3B42, JRA-55, and MERRA-2 for LT_d; and NCEP1 for LT_n), annual Bs for the remaining products were negative in <50% of grids, and even TRMM3B42RT, ERA-Interim, and NCEP2 showed negative Bs in <35% of grids. As shown in Figure 4(a2), most products had negative spring Bs for LT_{wd} and LT_d in >50% of grids, especially for PERSIANN, ERA-Interim, and MERRA-2 with >65% of grids. However, six of the 12 products underestimated LT_n in around 50% of grids in spring, followed by the other six products with overestimations in >50% of grids. Similar to spring, >50% of grids with negative Bs for summer LT_{wd} and LT_d were detected by most of the products, of which PERSIANN, GSMaP-RNL, ERA-Interim, and NCEP2 corresponded to >65% of grids (Figure 4(a3)). Except for PERSIANN, summer LT_n was underestimated in <50% of grids by the products, particularly for JRA-55, ERA-Interim, and ERA-5 with a grid percentage < 25%. In autumn (Figure 4(a4)), despite several exceptions, >50% of grids had negative Bs for the three LTs, and the PERSIANN product had negative Bs in >70% of grids. Relative to autumn cases, the opposite happened during winter (Figure 4(a5)), i.e., percentages of grids with

negative Bs for LT$_{wd}$, LT$_d$, and LT$_n$ being generally <50%, in particular for JRA-55 and ERA-Interim. In addition, based on percentages of grids with negative Bs for LT$_d$ and LT$_n$ (Figure 4(a1–a5)), differences generally exceeding 10% were identified on both annual and seasonal scales for most of products, particularly MERRA-2 and ERA-Interim with annual and summer differences around 40%, respectively. This suggests that, in terms of grid percentages corresponding to underestimated and overestimated precipitation LTs, the products' performance varies at daytime and nighttime.
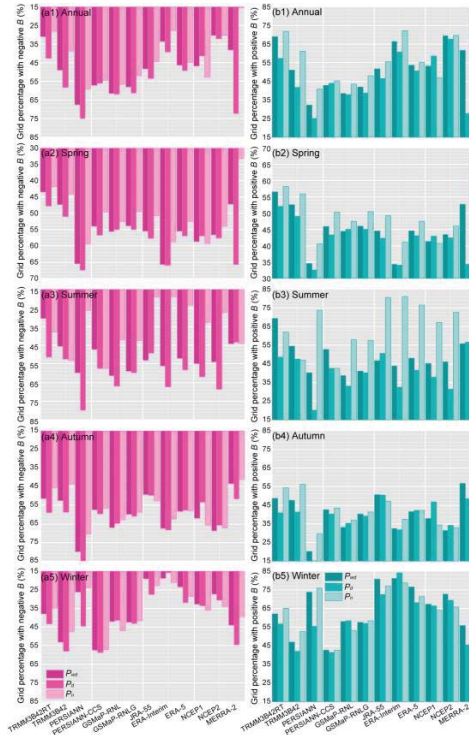
Figure 4. Grid percentages with negative (**a1–a5**) and positive biases (Bs; **b1–b5**) for annual and seasonal LTs across MC.

Taking MC as a whole, regional mean Bs for LT$_{wd}$, LT$_d$ and LT$_n$ derived from each product were calculated and are shown in Figure 5(a1–a5). At the annual scale, five products exhibited positive Bs for LT$_{wd}$, with a range from 0.39 mm/yr for TRMM3B42 to 10.10 mm/yr for NCEP2, while four products exhibited positive Bs for LT$_d$, ranging from 1.44 mm/yr for TRMM3B42RT to 5.01 mm/yr for NCEP2. Negative Bs were found for the remaining products, of which the lowest values of –7.88 mm/yr and –4.98mm/yr for LT$_{wd}$ and LT$_d$, respectively, were recorded for PERSIANN (Figure 5(a1)). In contrast, seven products overestimated annual LT$_n$, particularly ERA-Interim, NCEP2, and MERR-2 with Bs > 5 mm/yr, while the other products' Bs were all negative and generally <−3 mm/yr. Regarding the spring LTs (Figure 5(a2)), TRMM3B42RT, TRMM3B42, and MERRA-2 had positive Bs < 1.40 mm/yr, except for LT$_d$. However, negative spring Bs were found in the remaining products, ranging from −2.60 mm/yr (−1.61 mm/yr) for JRA-55 to −0.88 mm/yr (−0.51 mm/yr) for ERA-Interim for LT$_{wd}$ (LT$_d$), and from −1.52 mm/yr for NCEP1 to −0.31 mm/yr for PERSIANN-CCS for LT$_n$. For summer LT$_{wd}$ and LT$_n$ (Figure 5(a3)), most of the products exhibited positive Bs, while LT$_d$ was generally underestimated by the products (excluding TRMM3B42RT and NCEP1). Despite that, summer Bs for LTs were generally from −2 mm/yr to 2 mm/yr, except for TRMM3B42RT LT$_{wd}$ and LT$_n$, and MERRA-2 LT$_n$ with Bs >

2 mm/yr, and PERSIANN-CCS LT_{wd} and LT_d with Bs < −2 mm/yr. In autumn (Figure 5(a4)), absolute values of Bs for LTs from TRMM3B42, TRMM3B42RT, and MERRA-2 were all < 0.60 mm/yr, but Bs were generally < −1 mm/yr for the remaining products, and even some were lower than −4 mm/yr (i.e., PERSIANN, and NCEP2 for LT_{wd}). In contrast, the majority of products overestimated winter LT_{wd}, LT_d, and LT_n, and Bs were generally < 3 mm/yr, with exceptions of JRA-55, ERA-Interim, and ERA-5 having Bs > 3 mm/yr (Figure 5(a5)). In terms of Bs for annual LT_d and LT_n (Figure 5(a1)), there were differences for some products, i.e., TRMM3B42RT, PERSIANN-CCS, and ERA-Interim with large differences > 2 mm/yr, and TRMM3B42, JRA-55, ERA-5, NCEP1, and MERRA-2 showing different sign (positive/negative). Evident differences in Bs existed for some products in each season (Figure 3(a2–a5)); in summer there were eight products with different signs of Bs and four products with large differences (around ±1 mm/yr) but the same sign.
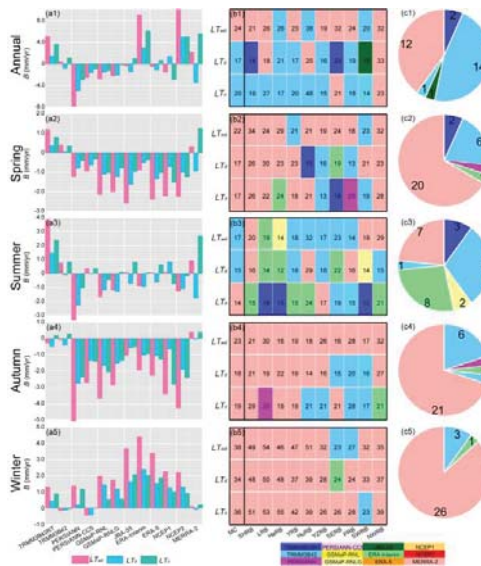
Figure 5. MC Bs derived from the selected 12 precipitation products (**a1–a5**), B-based optimal products (OPs) for MC and ten WRRs (**b1–b5**), and number of cases corresponding to B-based OPs on an annual or seasonal scale for ten WRRs (**c1–c5**). In figures (**b1–b5**), the number of each box represents grid percentage (%) of OP, which has been labelled with different colors. The number of figures (**c1–c5**) indicates the amount of a certain OP.

Considering offset effects of positive and negative Bs within MC and each WRR, we calculated the percentage of grids with the minimum absolute B for each product, and B-based OPs were identified as the product with the largest grid percentage (Figure 5(b1–b5,c1–c5)). Except for annual LT_d and LT_n and summer LT_{wd} and LT_d, for which the OP was TRMM3B42, the OP for all other LTs was MERRA-2 for MC (Figure 5(b1–b5)). For annual cases of the ten WRRs (Figure 5(c1)), the OPs were TRMM3B42 and MERRA-2 in 14 and 12 cases, respectively. Furthermore, the OP for most WRRs was MERRA-2 for annual LT_{wd}, and TRMM3B42 for both annual LT_d and LT_n (Figure 5(b1)). For B-based OPs of spring and autumn LT_{wd} (Figure 5(b2,b4)), the OP was MERRA-2 in most WRRs, while the OPs for LT_d and LT_n in southern and northern WRRs differed and were mainly TRMM3B42RT and MERRA-2, respectively. In summer (Figure 5(b3)), there were differences in OP for the three LTs, i.e., most WRRs with OPs of TRMM3B42 and MERRA-2 for LT_{wd}, TRMM3B42, MERRA-2, and ERA-Interim for LT_d, and ERA-Interim and TRMM3B42RT for LT_n. During winter (Figure 5(b5)), with four exceptions, all cases had an OP of MERRA-2. Overall, more than 20 cases had MERRA-2 as their OP for spring,

autumn, and winter, while TRMM3B42 had fewer than six cases. For summer, the OP was TRMM3B42 for nine cases, ERA-Interim for eight cases, and MERRA-2 for seven cases (Figure 5(c2–c5)).

3.4. Evaluation Using Error Metric

The MC RMSEs for LT_{wd}, LT_d and LT_n of each product are illustrated in Figure 6(a1–a5). For MC, TRMM3B42, TRMM3B42RT, PERSIANN, PERSIANN-CCS, and MERRA-2, RMSEs for annual LT_{wd}, LT_d and LT_n were lowest (<20.00 mm/yr for LT_{wd}; <10.00 mm/yr for LT_d and LT_n); this indicates that the accuracy of the five products, especially TRMM3B42, in detecting annual LTs is better. Except for PERSIANN-CCS and MERRA-2, the slightly smaller differences in annual RMSE for LT_d and LT_n from the remaining products suggests a comparable accuracy at daytime and nighttime. For LT_{wd}, LT_d, and LT_n, the largest RMSEs for each product occurred in summer and the smallest occurred in winter, due to their larger and smaller portion of annual MC precipitation, respectively. In each season, the minimum RMSE of the three LTs generally came from TRMM3B42 and MERRA-2, however, larger RMSEs were frequently found for GSMaP-RNL, GSMaP-RNLG, JRA-55, ERA-5, NCEP1, and NCEP2. Comparing RMSEs of LT_d and LT_n for each product in each season, LT_d for most of the products exhibited larger and smaller values in summer and the other three seasons, respectively; it should be noted that differences between LT_d and LT_n were not evident, excluding PERSIANN-CCS, which had an absolute difference > 2 mm/yr in winter.

Figure 6. MC root mean square error (RMSE) derived from the selected 12 precipitation products (**a1–a5**), RMSE-based optimal products (OPs) for MC and ten WRRs (**b1–b5**), and number of cases corresponding to RMSE-based OPs for annual or seasonal scale in ten WRRs (**c1–c5**). In figures (**b1–b5**), the number of each box represents RMSEs (mm/yr) of OP, which are labelled with different colors. The number of figures (**c1–c5**) indicates the amount of a certain OP.

Figure 6(b1–b5) illustrate the RMSE-based OPs of LTs for MC and WRRs. In general, the MC RMSE-based Ops for the three LTs for annual, spring (excluding LT_{wd}) and summer were TRMM3B42, while MERRA-2 was the autumn OP (excluding LT_n) and winter OP. At the annual scale, 13 cases had an RMSE-based OP of TRMM3B42, generally in southern WRRs, YRB, and HuRB; the remaining four WRRs had MERRA-2 as their OP (Figure 6(b1,c1)). In spring, the RMSE-based OP for the three LTs in northern WRRs and LT_{wd} in southern WRRs was MERRA-2, corresponding to 21 cases; eight cases had TRMM3B42 and PERSIANN as their OP, and these mainly appeared in LT_d and LT_n of southern WRRs (Figure 6(b2,c2)). During summer (Figure 6(b3,c3)), with three exceptions, all cases had TRMM3B42 (19 cases) and MERRA-2 (eight cases) as their OPs. In autumn, TRMM3B42 was the OP in seven cases mainly in southern WRRs (except for LT_{wd} and LT_d in YZRB and SERB), followed by MERRA-2, which was the OP in 21 cases (Figure 6(b4,c4)). With the exceptions of LT_{wd} and LT_d in YZRB and LT_{wd} in SWRB, MERRA-2 was the OP in 27 cases in winter (Figure 6(b5,c5)).

3.5. Evaluation Using Metric of Sign Accuracy

To examine the degree of agreement between the positive or negative sign of LTs from the products and the observed values, metrics of AS and JAS were computed over MC and are illustrated in Figure 7(a1–a5) and Figure 8(a1–a5), respectively. At the annual scale, MC AS values for LT_{wd}, LT_d, and LT_n from all products were > 50%. This suggests that the observed signs of LTs can be captured by the products, among which TRMM3B42RT, TRMM3B42, and MERRA-2 showed AS values > 70% for the three LTs, followed by PERSIANN, PERSIANN-CCS, ERA-Interim, and NCEP2 with values > 60% (Figure 7(a1)). During each season (Figure 7(a2–a5)), TRMM3B42RT, TRMM3B42, PERSIANN, PERSIANN-CCS, ERA-Interim, and MERRA-2 showed MC AS values > 60% for the three LTs, and the largest percentage (>70%) was found for TRMM3B42RT (except in winter), TRMM3B42 (except in winter), and MERRA-2. For the remaining six products, their AS-based performances differed among seasons. For example, all of them corresponded to autumn AS values > 50% for the three LTs; however, the values in the other seasons were generally < 50%. As shown in Figure 7(b1–b5), MC annual and summer AS-based OPs were MERRA-2 for LT_{wd}, but TRMM3B42 for LT_d and LT_n. For MC LTs in the remaining three seasons, the AS-based OP was MERRA-2, except for spring LT_d. Of the 30 annual cases in ten WRRs, AS-based OPs were MERRA-2 in 13 cases, TRMM3B42 in six cases and TRMM3B42RT in three cases, and there was more than one OP in five cases (Figure 7(b1–c1)). Among the ten WRRs, there were five or more OPs for each of the three LTs, indicating obvious regional differences for the products in detecting the same signs of LTs. In spring (Figure 7(b2–c2)), the AS-based OPs were MERRA-2 (15 cases), TRMM3B42 (five cases), TRMM3B42RT (four cases), ERA-Interim (four cases), and PERSIANN (two cases). Southern WRRs generally had OPs of TRMM3B42, TRMM3B42RT, and PERSIANN, while the OPs for northern WRRs were MERRA-2 and ERA-Interim. During summer (Figure 7(b3–c3)), AS-based OP was TRMM3B42 in 14 cases, mainly in southern WRRs, HaRB, YRB, and HuRB; and the OP was MERRA-2 in nine cases primarily in SHRB, LRB and NWRB. Of the 30 cases in autumn (Figure 7(b4–c4)), MERRA-2 was identified as the AS-based OP in 14 cases mainly in northern WRRs (excluding YRB), however, in eight cases the OP was TRMM3B42 generally in YRB, YZRB, and SWRB. Each of the six autumn cases in SERB and PRB had more than one OP. Regarding the 30 cases in winter (Figure 7(b5–c5)), 26 cases had AS-based OPs of MERRA-2 (22 cases) and TRMM3B42RT (four cases).
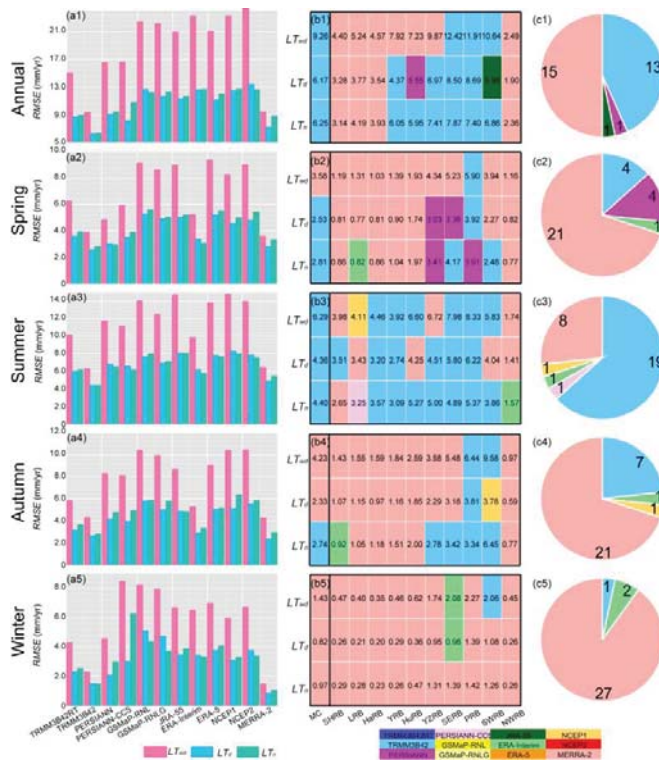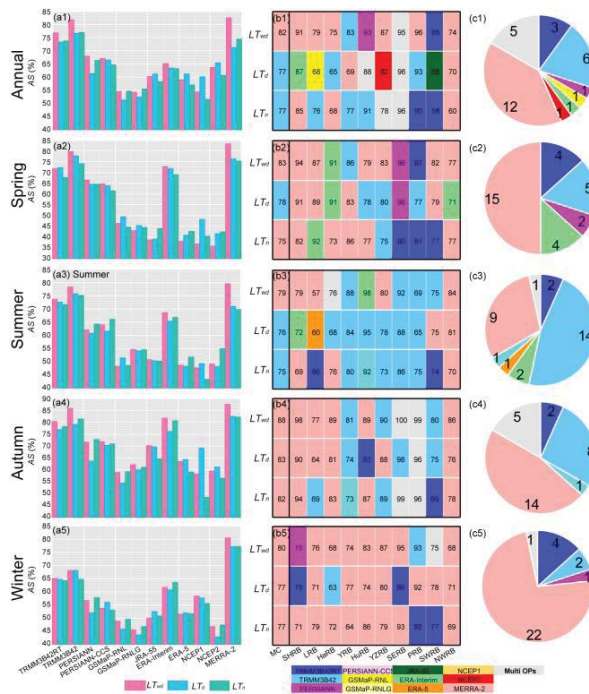
Figure 7. MC accurate sign (AS) values derived from the selected 12 precipitation products (**a1**–**a5**), AS-based optimal products (OPs) for MC and ten WRRs (**b1**–**b5**), and the number of cases corresponding to AS-based OPs (**c1**–**c5**) for the annual or seasonal scales in ten WRRs. For figures (**a1**–**a5**), AS is computed with Equation (5), indicating the degree of agreement between the positive or negative sign of precipitation trends from the products and the observed data. In figures (**b1**–**b5**), the number in each box represents AS values (%) of the OP, which have been labelled with different colors. The number of figures (**c1**–**c5**) indicates the amount of a certain OP.

Except for TRMM3B42RT, TRMM3B42, and MERRA-2 with annual JAS values > 55% (Figure 8(a1)), values of this metric were all below 50%, suggesting that these products have limited capacity to detect the co-variations of daytime and nighttime precipitation, in spite of relatively large AS for LTs (Figure 7(a1)). In spring (Figure 8(a2)), the best JAS-based performance was found in TRMM3B42 and MERRA-2 (with JAS around 60%), followed by TRMM3B42RT and ERA-Interim. Excluding PERSIANN and PERSIANN-CCS, the other six products had spring JAS values < 25%. JRA-55 and ERA-5 had JAS < 20%, which indicated that those six products could not capture the co-variations of spring precipitation changes at daytime and nighttime. During summer (Figure 8(a3)), TRMM3B42RT, TRMM3B42, and MERRA-2 performed the best (with JAS values > 53%), followed by PERSIANN, PERSIANN-CCS, and ERA-Interim with the next best performance (with JAS around 45%), and the remaining products (with JAS around 30%). For autumn (Figure 8(a4)), seven of the products correctly detected the co-variations of daytime and nighttime precipitation changes in >50% grids (i.e., JAS > 50%), particularly for TRMM3B42RT, TRMM3B42, ERA-Interim, and MERRA-2, which had JAS values > 64%. GSMaP-RNL and NCEP1 had JAS values near to 40% and performed the worst. Regarding winter JAS (Figure 8(a5)), values > 50% only appeared for MERRA-2, and the minima (around 25%) were found in GSMaP-RNL, GSMaP-RNLG, and NCEP2. As depicted in Figure 8b, MC annual and summer JAS-based OP was TRMM3B42, but for the other seasons the OP was MERRA-2. Except for SHRB and NWRB, with MERRA-2 as their annual JAS-based OP, the remaining WRRs generally had TRMM3B42 as the OP (Figure 8b). In southern WRRs, most had OPs of TRMM3B42 and

PERSIANN in spring and summer but MERRA-2 and PERSIANN in autumn and winter. By contrast, summer JAS-based OPs were MERRA-2 and TRMM3B42 in northern WRRs, while MERRA-2 was the OP in most northern WRRs.

Figure 8. MC joint AS (JAS) values derived from the selected 12 precipitation products (**a1–a5**), and JAS-based optimal products (OPs) for MC and ten WRRs (**b**). For figures (**a1–a5**), JAS is computed with Equation (6), indicating the capacity of a given product to rightly detect the signs of both P_d and P_n changes relative to the observed data. In (**b**), the number in each box represents JAS values (%) of the OP, which has been labelled with different colors.

4. Discussion

4.1. Possible Causes for Variation in Performance among Precipitation Products

In this study, we explored the reliability of the satellite-based and reanalysis products in capturing precipitation linear trends across MC, and found that the performances of these products exhibited clear differences. In general, TRMM3B42 and MERRA-2 showed the best overall performance. There are several possible explanations for the performance variation, e.g., input data, onboard sensors, and retrieval algorithm for the satellite-based products; and numerical models and their structures, parameterizations (especially for schemes about precipitation processes), and assimilation systems for the reanalysis products. Nonetheless, quantitatively identifying the impacts of these factors is difficult and beyond the scope of this study. As a result, we would like to discuss the potential causes of different performance among the precipitation products with the same retrieval algorithm or model structures, i.e., TRMM3B42RT vs. TRMM3B42, PERSIANN vs. PERSIANN-CCS, GSMaP-RNL vs. GSMaP-RNLG, ERA-Interim vs. ERA-5, and NCEP1 vs. NCEP2. It is evident that TRMM3B42 generally outperforms TRMM3B42RT, could be attributed to the fact that the former incorporates rain gauge data (i.e., monthly

GPCP and CAMS data; [32]) to adjust the precipitation estimates. In some WRRs, TRMM3B42RT performed better or was the OP, implying that the TRMM3B42 precipitation trends were occasionally overcorrected due to an inappropriate correction method (e.g., daily TRMM3B42RT adjusted with monthly observation; [32]). For PERSIANN and PERSIANN-CCS, their major differences are that the latter includes a cloud classification system based on cloud height, areal extent, and variability of texture estimated from satellite imagery to more accurately describe the relationship between precipitation rate and brightness temperature [35]. Despite that, the performance of PERSIANN in detecting precipitation trends was better than PERSIANN-CCS across MC based on most of the validation metrics. This indicates that the cloud classification system within PERSIANN-CCS has limited effectiveness in improving the estimated precipitation trends, although PERSIANN-CCS has been found to outperform PERSIANN in estimating precipitation amount over some regions of MC and its sub-regions (e.g., Tibetan Plateau and Yangtze River Basin, [88–91]). For the bias metric, PERSIANN-CCS performed better, mainly because, within a given region, the functions between precipitation rate and brightness temperature are established for each categorization of cloud-patch, and thus the regional biases are more likely to be offset. Relative to other satellite-based products, the two GSMaP products had the worst performance in MC and ten WRRs, indicating that the algorithm employed by GSMaP-RNL and GSMaP-RNLG may be problematic in capturing precipitation trends. Meanwhile, some studies also found that the GSMaP products had very low performance in capturing precipitation magnitudes and hydrological modeling over MC [92] and some Asian regions, such as the VuGia–ThuBon River Basin of Vietnam [93], and Mekong River Basin [94]. Moreover, the worst performance of GSMaP-RNLG in terms of specific validation metric suggests that its gauge-based correction processes are not efficient to adjust the precipitation trend. Relative to ERA-Interim, ERA-5 had a more advanced assimilation system and more and newer observational inputs, and thus was observed to have better performances (e.g., lower bias and root-mean squared error, and higher correlation coefficient) to reproduce precipitation in some regions, [45,91,95]. However, we found that in the study the ERA-5 precipitation trends poorly match the observations relative to the ERA-Interim. These findings are consistent with the findings of Nogueira across the globe [96], who pointed out that the trend of global-mean rainfall in ERA-Interim was closer to GPCP than ERA-5, and suggested that the possible causes were associated with the global energy budget [97,98]. Due to NCEP2 with new system components including simple precipitation assimilation over land surfaces for improved soil wetness [43], NCEP2 precipitation agreed more closely with gauge measurements than NCEP1 data in China [99], USA [100], and Central Equatorial Africa [101]; by contrast, comparisons between NCEP1 and NCEP2 in representing precipitation linear trends show that no obvious differences existed. This may be related to significant time-varying jumps in the late 2000s within NCEP2, mainly due to the changes in observing systems, such as the introduction of new data into the assimilation systems [102,103]; this is also the possible cause of poor performance for JRA-55 [101]. The validation metrics clearly show that, based on precipitation linear trends, MERRA-2 performed better than other reanalysis products and even satellite-based precipitation products in MC. The better performances of MEERA-2 for representing precipitation amount were also found in other regions (e.g., Nepal, and the Pamir region of Tajikistan, [104,105]). Some scholars pointed that the possible causes are related to the advanced data assimilation technique within MERRA-2 and the bias corrections of MERRA-2 precipitation [45]. We should note that for a given product there are differences in performance of detecting precipitation trends within a day (e.g., daytime and nighttime annual correlation coefficients for ERA-Interim) and among seasons (e.g., smaller winter correlation coefficient values for PERSIANN but larger values in the other three seasons), mainly due to the different physical mechanisms controlling precipitation processes [69–71,73,74]. For example, some studies stated that sea–land breeze is closely associated with the diurnal cycle of precipitation in coastal areas, while topography and mountain–valley breeze plays an important role in the interior [73,74]. Therefore, to increase the accuracy of sub-daily and seasonal precipitation estimates, specific algorithms for the satellite-based products and specific model structures for the reanalysis products should be developed.

4.2. Uncertainties from Rain Gauge Data

We employed rain gauge data as a reference to validate the 12 precipitation products in detecting precipitation trends at different time scales across MC. It should be noted that the inherent uncertainties within the gauge data, which are related to flaws in calibration, wind-related under-catch, and wetting and evaporation losses, could bias the gauge measurements from the real values and weaken the robustness of the validation results (e.g., [106–109]). For example, Shedekar et al. [109] found that relative to the actual rainfall depths, the precipitation measurements from three calibrated tipping-bucket rain gauges were underestimations, particularly for heavy rainfall, and they highlighted that the biases were closely associated with the gauge calibrations. When it is windy, gauge observations are often impacted by wind-related under-catch effects through deflecting the flow and inducing eddies and turbulence around the gauges [108–111]. In general, wind can cause some raindrops, especially smaller ones, to miss the funnel or fall at an inclination, and finally impact the catch efficiency of the gauges. To what extent wind influences the accuracy of the gauge measurements is dependent on ambient wind speed, raindrop size distribution, and gauge design [110]. Sieck et al. [110] reported that, compared to rainfall from collocated buried gauges, wind-exposed aboveground gauges would likely observe about 2–10% less precipitation. Due to water adhering to the inside walls of the gauge and then evaporating, the gauge-recorded precipitation is generally lower than the true value, and the biases vary among gauge configurations (e.g., frequency of emptying) and precipitation types [106,112,113]. A Russian study revealed that, for each record of rainfall measurement, the mean average wetting loss was 0.2 mm, but for both snow and mixed precipitation the value was 0.15 mm [112]. Due to being exposed to the atmosphere, water within rain gauges is usually evaporated (i.e., evaporation losses; [114–117]). It is reported that evaporation losses for gauged precipitation generally range from 0.1 to 0.8 mm/day or from 0 to 1%; however, the magnitudes differ among gauge types, climate backgrounds, and seasons [114]. The combined effects induced by the aforementioned factors on rain gauge measurements are likely to underestimate the recorded precipitation [118], e.g., the bias-corrected annual precipitation (removing the uncertainties within raw observations) being 30–330 mm or 10–65% higher than the raw observations over Siberia.

Usually, the quality of precipitation observations is accompanied by an issue of standardization, or lack of, which is mainly due to changes in gauge instruments, station relocation and environment, etc. [119–122]. Moreover, these factors result in negative impacts on data quality, in particular for climate researches using long-term time series (e.g., linear trend evaluation in this study). Before using the gauged precipitation measurements, it is necessary to reduce and even remove the associated uncertainties, e.g., adjust the raw records using metadata about gauges, and at least eliminating sites with non-homogeneous measurements identified by some statistic methods. The Pettitt test has been used to remove sites with non-homogenous measurements due to a lack of metadata for the selected sites, but there is no guarantee that the remaining sites have no issues, which can weaken the confidence level of the results. Besides, mismatches between representatives of gauge precipitation (i.e., a point of space in time accumulation) and selected products (i.e., a snapshot of time in space aggregation) are likely to have an effect on the accuracy and precision of qualitative and quantitative assessments of various precipitation products [123–125]. For instance, the spatial resolution of all the 12 products is generally lower than $0.25° \times 0.25°$ (except for PERSIANN-CCS), across which the estimated precipitation was averaged, while the spatial representation of a gauge is much smaller than the coverage of the pixel of the 12 products. Considering the variability of precipitation over a small spatial extent, a sparse gauge network may not identify meso-/micro-scale weather system-associated precipitation (e.g., convective precipitation; [126–129]); thus, gauge precipitation measurements may be smaller in magnitude and frequency than the ground-truthed values for a given pixel.

5. Conclusions

As important surrogate for precipitation estimates, various satellite-based and reanalysis precipitation products need to be validated from different perspectives. Especially, the information

about the capacity of the satellite-based and reanalysis precipitation is scarce on a sub-daily scale, especially for China. However, the assessments regarding precipitation trends are fundamental for selecting the reliable products to explore precipitation changes, particularly for regions with limited or even no observations. Thus, with a motivation to explore twelve popular precipitation products (i.e., six satellite-based and six reanalysis products) in detecting precipitation linear trends across MC, we collected daytime and nighttime observations from a dense rain gauge network during 2003–2017, and examined LT_{wd}, LT_d, and LT_n across mainland China. We found that annual and seasonal LT_{wd}, LT_d, and LT_n for MC and most WRRs were positive but with regional differences. In terms of magnitude and sign (i.e., decreasing and increasing), LT_d and LT_n in a certain region showed evident differences, confirming the necessity to evaluate precipitation products at a sub-daily scale. Then, several statistical metrics (i.e., CC, B, RMSE, AS, and JAS) were employed to identify the differences and agreements of LTs for MC and ten WRRs between twelve precipitation products and gauge observations on sub-daily scale. In general, values of a given metric for annual and seasonal LT_{wd}, LT_d, or LT_n differed among products. Meanwhile, performances for single product varied among seasons and between daytime and nighttime. At last, the metric-based OPs were identified for MC and each WRR. The metric-based OPs varied among regions and seasons, and between daytime and nighttime, but the most frequent OPs were TRMM3B42, ERA-Interim, and MERRA-2.

The comparison of satellite-based and reanalysis products in ability to detect precipitation linear trends in this study provides suggestions for developers and the potential users of these products across mainland China. For a given product, varying performance for different validation metrics at different timescales (between daytime and nighttime) suggests that the product's group can try to develop specific algorithms/models during a certain season (at a sub-daily scale) and correction procedures to improve its capacity to reproduce precipitation trends. For the potential users who focus on long-term precipitation changes across MC, this study provides necessary and detailed information about the existing popular precipitation products' performances in detecting linear trends, which is fundamental to obtaining robust conclusions.

References

1. Cui, Y.; Yan, D.; Hong, T.; Xiao, C.; Luo, X.; Zhang, Q. Comparison of typical year and multiyear building simulations using a 55-year actual weather data set from China. *Appl. Energy* **2017**, *195*, 890–904. [CrossRef]
2. Kidd, C.; Becker, A.; Huffman, G.J.; Muller, C.L.; Joe, P.; Skofronick-Jackson, G.; Kirschbaum, D.B. So How much of the Earth's surface is covered by rain gauges? *Bull. Am. Meteorol. Soc.* **2017**, *98*, 69–78. [CrossRef] [PubMed]
3. Zhou, C.; Wang, K. Contrasting daytime and nighttime precipitation variability between observations and eight reanalysis products from 1979 to 2014 in China. *J. Clim.* **2017**, *30*, 6443–6464. [CrossRef]

4. Milly, P.C.D.; Dunne, K.A.; Vecchia, A.V. Global patterns of trends in streamflow and water availability in a changing climate. *Nature* **2005**, *438*, 347–350. [CrossRef] [PubMed]

5. Sorooshian, S.; AghaKouchak, A.; Arkin, P.; Eylander, J.; Foufoula-Georgiou, E.; Harmon, R.; Hendrickx, J.M.H.; Imam, B.; Kuligowski, R.; Skahill, B.; et al. Advanced concepts on remote sensing of precipitation at multiple scales. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 1353–1357. [CrossRef]

6. Sun, S.; Zhou, S.; Shen, H.; Chai, R.; Chen, H.; Liu, Y.; Shi, W.; Wang, J.; Wang, G.; Zhou, Y. Dissecting Performances of PERSIANN-CDR Precipitation Product over Huai River Basin, China. *Remote Sens.* **2019**, *11*, 1805. [CrossRef]

7. Thornes, J.; Bloss, W.; Bouzarovski, S.; Cai, X.; Chapman, L.; Clark, J.; Dessai, S.; Du, S.; van der Horst, D.; Kendall, M.; et al. Communicating the value of atmospheric services. *Meteorol. Appl.* **2010**, *17*, 243–350. [CrossRef]

8. Murthy, C.S.; Singh, J.; Kumar, P.; Sai, M.V.R.S. A composite index for drought hazard assessment using CPC rainfall time series data. *Int. J. Environ. Sci. Technol.* **2017**, *14*, 1981–1988. [CrossRef]

9. Sun, S.; Chen, H.; Ju, W.; Wang, G.; Sun, G.; Huang, J.; Ma, H.; Gao, C.; Hua, W.; Yan, G. On the coupling between precipitation and potential evapotranspiration: Contributions to decadal drought anomalies in the Southwest China. *Clim. Dyn.* **2017**, *48*, 3779–3797. [CrossRef]

10. Sun, S.; Li, Q.; Li, J.; Wang, G.; Zhou, S.; Chai, R.; Hua, W.; Deng, P.; Wang, J.; Lou, W. Revisiting the evolution of the 2009–2011 meteorological drought over Southwest China. *J. Hydrol.* **2019**, *568*, 385–402. [CrossRef]

11. Wei, J.; Jin, Q.; Yang, Z.-L.; Dirmeyer, P.A. Role of ocean evaporation in California droughts and floods. *Geophys. Res. Lett.* **2016**, *43*, 6554–6562. [CrossRef]

12. Berghuijs, W.R.; Aalbers, E.E.; Larsen, J.R.; Trancoso, R.; Woods, R.A. Recent changes in extreme floods across multiple continents. *Environ. Res. Lett.* **2017**, *12*. [CrossRef]

13. 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]

14. Nguyen, P.; Ombadi, M.; Sorooshian, S.; Hsu, K.; AghaKouchak, A.; Braithwaite, D.; Ashouri, H.; Thorstensen, A.R. The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 5801–5816. [CrossRef]

15. Pan, M.; Li, H.; Wood, E.F. Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resour. Res.* **2010**, *46*, W09535. [CrossRef]

16. Xie, P.; Janowiak, J.E.; Arkin, P.A.; Adler, R.; Gruber, A.; Ferraro, R.; Huffman, G.J.; Curtis, S. GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Clim.* **2003**, *16*, 2197–2214. [CrossRef]

17. Kidd, C.; Kniveton, D.R.; Todd, M.C.; Bellerby, T.J. Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeorol.* **2003**, *4*, 1088–1104. [CrossRef]

18. Wentz, F.J.; Ricciardulli, L.; Hilburn, K.; Mears, C. How much more rain will global warming bring? *Science* **2007**, *317*, 233–235. [CrossRef]

19. Ebert, E.E.; Manton, M.J.; Arkin, P.A.; Allam, R.J.; Holpin, C.E.; Gruber, A. Results from the GPCP Algorithm Intercomparison Programme. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 2875–2887. [CrossRef]

20. Sapiano, M.R.P.; Arkin, P.A. An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeorol.* **2009**, *10*, 149–166. [CrossRef]

21. Gou, Y.; Ma, Y.; Chen, H.; Weng, Y. Radar-derived quantitative precipitation estimation in complex terrain over the eastern Tibetan Plateau. *Atmos. Res.* **2018**, *203*, 286–297. [CrossRef]

22. Maddox, R.A.; Zhang, J.; Gourley, J.J.; Howard, K.W. Weather radar coverage over the contiguous United States. *Weather Forecast.* **2002**, *17*, 927–934. [CrossRef]

23. Turk, F.J.; Haddad, Z.S.; Kirstetter, P.; You, Y.; Ringerud, S. An observationally based method for stratifying a priori passive microwave observations in a Bayesian-based precipitation retrieval framework. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 145–164. [CrossRef]

24. Kuligowski, R.J.; Li, Y.; Zhang, Y. Impact of TRMM data on a low-latency, high-resolution precipitation algorithm for flash-flood forecasting. *J. Appl. Meteorol. Climatol.* **2013**, *52*, 1379–1393. [CrossRef]

25. Ashouri, H.; Hsu, K.; Sorooshian, S.; Braithwaite, D.K.; Knapp, K.R.; Cecil, L.D.; Nelson, B.R.; Prat, O.P. PERSIANNCDR: Daily precipitation climate data record from multi-satellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 69–83. [CrossRef]

26. Adler, R.F.; Huffman, G.J.; Chang, A.; Ferraro, R.; Xie, P.P.; Janowiak, J.; Rudolf, B.; Scjneider, U.; Curtis, S.; Gruber, D.B.A.; et al. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **2003**, *4*, 1147–1167. [CrossRef]

27. Gebregiorgis, A.; Hossain, F. How much can a priori hydrologic model predictability help in optimal merging of satellite precipitation products? *J. Hydrometeorol.* **2011**, *12*, 1287–1298. [CrossRef]

28. Michaelides, S.; Levizzani, V.; Anagnostou, E.; Bauer, P.; Kasparis, T.; Lane, J.E. Precipitation: Measurement, Remote Sensing, Climatology and Modeling. *Atmos. Res.* **2009**, *94*, 512–533. [CrossRef]

29. Prigent, C. Precipitation retrieval from space: An overview. *Comptes Rendus Geosci.* **2010**, *342*, 380–389. [CrossRef]

30. Screen, J.A.; Simmonds, I. Erroneous arctic temperature trends in the ERA-40 reanalysis: A closer look. *J. Clim.* **2011**, *24*, 2620–2627. [CrossRef]

31. Turk, J.T.; Mostovoy, G.V.; Anantharaj, V. The NRL-Blend High Resolution Precipitation Product and its Application to Land Surface Hydrology. In *Satellite Rainfall Applications for Surface Hydrology*; Gebremichael, M., Hossain, F., Eds.; Springer: Dordrecht, The Netherlands, 2010.

32. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.J.; Bowman, K.P.; Yong, Y.; Stocker, E.F.; Wolff, D.B. The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-global, multi-year, combined-sensor precipitation at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]

33. Hsu, K.L.; Gao, X.G.; Sorooshian, S.; Gupta, H.V. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.* **1997**, *36*, 1176–1190. [CrossRef]

34. Sorooshian, S.; Hsu, K.-L.; Gao, X.; Gupta, H.V.; Imam, B.; Braithwaite, D. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2035–2046. [CrossRef]

35. Hong, Y.; Hsu, K.L.; Sorooshian, S.; Gao, X. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteorol.* **2004**, *43*, 1834–1853. [CrossRef]

36. Ushio, T.; Kachi, M. Kalman Filtering Applications for Global Satellite Mapping of Precipitation (GSMaP). In *Satellite Rainfall Applications for Surface Hydrology*; Gebremichael, M., Hossain, F., Eds.; Springer: Dordrecht, The Netherlands, 2010.

37. 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]

38. Huffman, G.J. README for Accessing Experimental Realtime TRMM Multi-Satellite Precipitation Analysis (TMPART) Data Sets. NASA Tech. Doc. 2015, 12. Available online: ftp://mesoa.gsfc.nasa.gov/pub/trmmdocs/rt/3B4XRT_README.pdf (accessed on 10 August 2019).

39. Funk, C.C.; Peterson, P.J.; Landsfeld, M.F.; Pedreros, D.H.; Verdin, J.P.; Sukla, S.; Husak, G.J.; Rowland, J.D.; Harrison, L.; Hoell, A.; et al. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 1–21. [CrossRef]

40. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.J.M.; Weedon, G.P.; Brocca, L.; Pappenberger, F.; Huffman, G.J.; Wood, E.F. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6201–6217. [CrossRef]

41. 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]

42. Kalnay, E. NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–472. [CrossRef]

43. Kanamitsu, M.; Ebisuzaki, W.; Woollen, J.; Yang, S.K.; Hnilo, J.J.; Fiorino, M.; Potter, G.L. NCEP–DOE AMIP-II reanalysis (R-2). *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1631–1643. [CrossRef]

44. Kobayashi, S.; Ota, Y.; Harada, Y.; Ebita, A.; Moriya, M.; Onoda, H.; Onogi, K.; Kamahori, H.; Kobayashi, C.; Endo, H.; et al. The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteorol. Soc. JPN* **2015**, *93*, 5–48. [CrossRef]

45. Molod, A.; Takacs, L.; Suarez, M.; Bacmeister, J. Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA-2. *Geosci. Model. Dev.* **2015**, *8*, 1339–1356. [CrossRef]

46. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]

47. Bengtsson, L.; Hodges, K.; Esch, M.; Keenlyside, N.; Kornblueh, L.; Luo, J.; Yamagata, T. How may tropical cyclones change in a warmer climate? *Tellus A* **2007**, *59*, 539–561. [CrossRef]

48. Chen, G.; Iwasaki, T.; Qin, H.; Sha, W. Evaluation of the warm-season diurnal variability over East Asia in recent reanalyses JRA-55, ERA-Interim, NCEP CFSR, and NASA MERRA. *J. Clim.* **2014**, *27*, 5517–5537. [CrossRef]

49. Saha, S.; Moorthi, S.; Pan, H.-L.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Kistler, R.; Wollen, J.; Berhringer, D.; et al. The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.* **2010**, *91*, 1015–1058. [CrossRef]

50. Szczypta, C.; Calvet, J.C.; Albergel, C.; Balsamo, G.; Boussetta, S.; Carrer, D.; Lafont, S.; Meurey, C. Verification of the new ECMWF ERA-Interim reanalysis over France. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 647–666. [CrossRef]

51. Anagnostou, E.N. Overview of overland satellite rainfall estimation for hydro-meteorological applications. *Surv. Geophys.* **2004**, *25*, 511–537. [CrossRef]

52. Hodges, K.I.; Lee, R.W.; Bengtsson, L. A comparison of extratropical cyclones in recent reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25. *J. Clim.* **2011**, *24*, 4888–4906. [CrossRef]

53. Laviola, S.; Levizzani, V.; Cattani, E.; Kidd, C. The 183-wsl fast rain rate retrieval algorithm. part II: Validation using ground radar measurements. *Atmos. Res.* **2013**, *134*, 77–86. [CrossRef]

54. Lin, R.; Zhou, T.; Qian, Y. Evaluation of global monsoon precipitation changes based on five reanalysis datasets. *J. Clim.* **2014**, *27*, 1271–1289. [CrossRef]

55. Maggioni, V.; Sapiano, M.R.P.; Adler, R.F. Estimating uncertainties in high-resolution satellite precipitation products: Systematic or random error? *J. Hydrometeorol.* **2016**, *17*, 1119–1129. [CrossRef]

56. Urraca, R.; Huld, T.; Gracia-Amillo, A.; Martinez-De-Pison, F.J.; Kaspar, F.; Sanz-Garcia, A. Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Sol. Energy* **2018**, *164*, 339–354. [CrossRef]

57. Zhao, D.M.; Fu, C.B. Comparisons of low-level circulation characteristics between ECHAM5/MPI-OM results and NCEP/NCAR re-analysis data in East Asia. *Atmos. Ocean. Sci. Lett.* **2010**, *3*, 189–194.

58. Sungmin, O.; Kirstetter, P.E. Evaluation of diurnal variation of GPM IMERG derived summer precipitation over the contiguous US using MRMS data. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 270–281.

59. Omranian, E.; Sharif, H.O. Evaluation of the Global Precipitation Measurement (GPM) satellite rainfall products over the Lower Colorado River Basin, Texas. *J. Am. Water Resour.* **2018**, 1–17. [CrossRef]

60. Mahto, S.S.; Mishra, V. Does ERA-5 Outperform other reanalysis products for hydrologic applications in India? *J. Geophys. Res. Atmos.* **2018**, *124*, 9423–9441. [CrossRef]

61. Li, C.; Tang, G.; Hong, Y. Cross-evaluation of ground-based, multi-satellite and reanalysis precipitation products: Applicability of the triple collocation method across Mainland China. *J. Hydrol.* **2018**, *562*, 71–83. [CrossRef]

62. Huang, D.-Q.; Zhu, J.; Zhang, Y.-C.; Huang, Y.; Kuang, X.-Y. Assessment of summer monsoon precipitation derived from five reanalysis datasets over East Asia. *Q. J. R. Meteorol. Soc.* **2016**, *142*, 108–119. [CrossRef]

63. Bai, P.; Liu, X. Evaluation of five satellite-based precipitation products in two gauge-scarce basins on the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 1316. [CrossRef]

64. Fan, B.; Luo, G.; Zhang, C.; Hu, Z.; Li, C.; Wang, Y.; Bai, L. Evaluation of summer precipitation of CFSR, ERA-Interim and MERRA reanalyses in Xinjiang. *Geophys. Res.* **2013**, *32*, 1602–1612, (In Chinese with English Abstract).

65. de Leeuw, J.; Methven, J.; Blackburn, M. Evaluation of ERA-Interim reanalysis precipitation products using England and Wales observations. *Q. J. R. Meteorol. Soc.* **2015**, *141*, 798–806. [CrossRef]

66. Derin, Y.; Anagnostou, E.N.; Berne, A.; Borga, M.; Boudevillain, B.; Buytaert, W.; Chang, C.-H.; Delrieu, G.; Hong, Y.; Hsu, Y.C.; et al. Multi-regional satellite precipitation products evaluation over complex terrain. *J. Hydrometeorol.* **2016**, *17*, 1817–1836. [CrossRef]

67. Guilloteau, C.; Roca, R.; Gosset, M. A Multiscale Evaluation of the Detection Capabilities of High-Resolution Satellite Precipitation Products in West Africa. *J. Hydrometeorol.* **2016**, *17*, 2041–2059. [CrossRef]

68. Lorenz, C.; Kunstmann, H. The hydrological in three state-of-the-art reanalysis: Intercomparison and performance analysis. *J. Hydrometeorol.* **2012**, *13*, 1397–1420. [CrossRef]

69. Yu, R.; Yuan, W.; Li, J. The asymmetry of rainfall process. *Chin. Sci. Bull.* **2013**, *58*, 1850–1856. [CrossRef]

70. Yu, R.; Yuan, W.; Li, J.; Fu, Y. Diurnal phase of late-night against late-afternoon of stratiform and convective precipitation in summer southern contiguous China. *Clim. Dyn.* **2010**, *35*, 567–576. [CrossRef]

71. Cheng, P.; Gao, L.; Zuo, X.; Zhong, F. Statistical analyses of spatial and temporal variabilities in total, daytime, and nighttime precipitation indices and of extreme dry/wet association with large-scale circulations of Southwest China, 1961–2016. *Atmos. Res.* **2019**, *219*, 166–182. [CrossRef]

72. Ghate, V.P.; Kollias, P. On the controls of daytime precipitation in the Amazonian dry season. *J. Hydrometeorol.* **2016**, *17*, 3079–3097. [CrossRef]

73. Wallace, J.M. Diurnal variations in precipitation and thunderstorm frequency over the conterminous united states. *Mon. Weather Rev.* **1975**, *103*, 406–419. [CrossRef]

74. Yu, R.; Xu, Y.; Zhou, T.; Li, J. Relation between rainfall duration and diurnal variation in the warm season precipitation over central eastern China. *Geophys. Res. Lett.* **2007**, *34*, L13703. [CrossRef]

75. Brown, K.; Kamruzzaman, M.; Beecham, S. Trends in sub-daily precipitation in Tasmania using regional dynamically downscaled climate projections. *J. Hydrol. Reg. Stud.* **2017**, *10*, 18–34. [CrossRef]

76. Lenderink, G.; van Meijgaard, E. Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environ. Res. Lett.* **2010**, *5*, 252–258. [CrossRef]

77. Lenderink, G.; Mok, H.Y.; Lee, T.C.; van Oldenborgh, G.J. Scaling and trends of hourly precipitation extremes in two different climate zones: Hong Kong and the Netherlands. *Hydrol. Earth Syst. Sci. Dis.* **2011**, *8*, 4701–4719. [CrossRef]

78. Liu, L.Y.; Ma, Z.-G. Intra-annual variability of diurnal cycle precipitation over china from 1960–2000. *Atmos. Ocean. Sci. Lett.* **2013**, *6*, 451–456.

79. Mao, J.Y.; Wu, G.X. Diurnal variations of summer precipitation over the Asian monsoon region as revealed by TRMM satellite data. *Sci. China Earth Sci.* **2012**, *55*, 554–566. [CrossRef]

80. Liu, X.; Zhang, M.; Wang, S.; Wang, J.; Zhao, P.; Zhou, P. Assessment of diurnal variation of summer precipitation over the Qilian Mountains based on an hourly merged dataset from 2008 to 2014. *J. Geogr. Sci.* **2017**, *27*, 326–336. [CrossRef]

81. Lin, C.; Liu, L.; Lin, W.; Bai, Y.; Qi, H.; Yang, H. Characteristics of summer precipitation diurnal variations in Hubei Province. *Trans. Atmos. Sci.* **2016**, *39*, 490–500, (In Chinese with English Abstract).

82. Han, Y.; Ma, Z.; Yang, Q.; Pan, Z. Changing Characteristics of Daytime and Nighttime Precipitation in Xinjiang under Global Warming. *Clim. Environ. Res.* **2014**, *19*, 763–772, (In Chinese with English Abstract).

83. Chen, F.; Gao, Y. Evaluation of precipitation trends from high-resolution satellite precipitation products over Mainland China. *Clim. Dyn.* **2018**, *51*, 3311–3331. [CrossRef]

84. Ren, Z.H.; Zhao, P.; Zhang, Q.; Zhang, Z.F.; Chen, Z. Quality control procedures for hourly precipitation data from automatic weather stations in China. *Meteorol. Mon.* **2010**, *36*, 123–132, (In Chinese with English Abstract).

85. The National Aeronautics and Space Administration (NASA) Shuttle Radar Topographic Mission (SRTM) 90m Digital Elevation Model (DEM) Digital Elevation Database. Available online: http://srtm.csi.cgiar.org/ (accessed on 7 July 2019).

86. Huang, R.; Chen, W.; Ding, Y.; Li, C. Studies on the monsoon dynamics and the interaction between monsoon and ENSO cycle. *Chin. J. Atmos. Sci.* **2003**, *27*, 484–502, (In Chinese with English Abstract).

87. Hsu, K.; Sellars, S.; Nguyen, P.; Braithwaite, D.; Chu, W. G-WADI PERSIANN-CCS GeoServer for extreme precipitation event monitoring. *Sci. Cold Arid Reg.* **2013**, *5*, 6–15.

88. Liu, Y.; Wu, Y.; Feng, Z.; Huang, X.; Wang, D. Evaluation of a Variety of Satellite Retrieved Precipitation Products Based on Extreme Rainfall in China. *Trop. Geogr.* **2007**, *37*, 417–433, (In Chinese with English Abstract).

89. Shen, Y.; Xiong, A.; Wang, Y.; Xie, P. Performance of high-resolution satellite precipitation products over China. *J. Geophys. Res. Atmos.* **2010**, *115*, D02114. [CrossRef]

90. Wang, Y.; Xie, X.; Meng, S.; Wu, D.; Chen, Y.; Jiang, F.; Zhu, B. Magnitude agreement, occurrence consistency, and elevation dependency of satellite-based precipitation products over the Tibetan Plateau. *Remote Sens.* **2020**, *12*, 1750. [CrossRef]

91. Xiao, S.; Xia, J.; Zou, L. Evaluation of multi-satellite precipitation products and their ability in capturing the characteristics of extreme climate events over the Yangtze River Basin, China. *Water* **2020**, *12*, 1179. [CrossRef]

92. Qin, Y.; Chen, Z.; Shen, Y.; Zhang, S.; Shi, R. Evaluation of Satellite Rainfall Estimates over the Chinese Mainland. *Remote Sens.* **2014**, *6*, 11649–11672. [CrossRef]

93. Ngo-Duc, T.; Matsumoto, J.; Kamimera, H.; Bui, H.-H. Monthly adjustment of Global Satellite Mapping of Precipitation (GSMaP) data over the VuGia–ThuBon River Basin in Central Vietnam using an artificial neural network. *Hydrol. Res. Lett.* **2013**, *7*, 85–90. [CrossRef]

94. Try, S.; Tanaka, S.; Tanaka, K.; Sayama, T.; Oeurng, C.; UK, S.; Takara, K.; Hu, M.; Han, D. Comparison of gridded precipitation datasets for rainfall-runoff and inundation modeling in the Mekong River Basin. *PLoS ONE* **2020**, *15*, e0226814. [CrossRef]

95. Fallah, A.; Rakhshandehroo, G.R.; Berg, P.; Sungmin, O.; Orth, R. Evaluation of precipitation datasets against local observations in southwestern Iran. *Int. J. Climatol.* **2020**, *40*, 4102–4116. [CrossRef]

96. Nogueira, M. Inter-comparison of ERA-5, ERA-Interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. *J. Hydrol.* **2020**, *583*, 124632. [CrossRef]

97. Nogueira, M. The multi-scale structure of atmospheric energetic constraints on globally averaged precipitation. *Earth Syst. Dyn.* **2019**, *10*, 219–232. [CrossRef]

98. Stephens, G.L.; Ellis, T.D. Controls of global-mean precipitation increases in global warming GCM experiments. *J. Clim.* **2008**, *21*, 6141–6155. [CrossRef]

99. Ma, L.; Zhang, T.; Frauenfeld, O.W.; Ye, B. Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China. *J. Geophys. Res. Atmos.* **2009**, *114*, D09105. [CrossRef]

100. Tang, D.; Ma, C.; Wang, Y.; Xu, X. Multiscale evaluation of NCEP and CRUNCEP data sets at 90 large U.S. cities. *J. Geophys. Res. Atmos.* **2017**, *122*, 7433–7444. [CrossRef]

101. Hua, W.; Zhou, L.; Nicholson, S.E.; Chen, H.; Qin, M. Assessing reanalysis data for understanding rainfall climatology and variability over Central Equatorial Africa. *Clim. Dyn.* **2019**, *53*, 5139. [CrossRef]

102. Trenberth, K.E.; Fasullo, J.T.; Mackaro, J. Atmospheric moisture transports from ocean to land and global energy flows in reanalyses. *J. Clim.* **2011**, *24*, 4907–4924. [CrossRef]

103. Zhang, L.; Kumar, A.; Wang, W. Influence of changes in observations on precipitation: A case study for the Climate Forecast System Reanalysis (CFSR). *J. Geophys. Res. Atmos.* **2012**, *117*, D08105. [CrossRef]

104. Hamal, K.; Sharma, S.; Khadka, N.; Baniya, B.; Ali, M.; Shrestha, M.S.; Xu, T.; Shrestha, D.; Dawadi, B. Evaluation of MERRA-2 Precipitation Products Using Gauge Observation in Nepal. *Hydrology* **2020**, *7*, 40. [CrossRef]

105. Zandler, H.; Haag, I.; Samimi, C. Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. *Sci. Rep.* **2019**, *9*, 1–15. [CrossRef] [PubMed]

106. Devine, K.A.; Mekis, É. Field accuracy of Canadian rain measurements. *Atmos. Ocean.* **2008**, *46*, 213–227. [CrossRef]

107. Lanza, L.G.; Vuerich, E. The WMO Field Intercomparison of Rain Intensity Gauges. *Atmos. Res.* **2009**, *94*, 534–543. [CrossRef]

108. Pollock, M.D.; O'Donnell, G.; Quinn, P.; Dutton, M.; Black, A.; Wilkinson, M.E.; Colli, M.; Stagnaro, M.; Lanza, L.G.; Lewis, E.; et al. Quantifying and mitigating wind-induced undercatch in rainfall measurements. *Water Resour. Res.* **2018**, *54*, 3863–3875. [CrossRef]

109. Shedekar, V.S.; King, K.W.; Fausey, N.R.; Soboyejo, A.B.O.; Harmel, R.D.; Brown, L.C. Assessment of measurement errors and dynamic calibration methods for three different tipping bucket rain gauges. *Atmos. Res.* **2016**, *178*, 445–458. [CrossRef]

110. Sieck, L.C.; Burges, S.J.; Steiner, M. Challenges in obtaining reliable measurements of point rainfall. *Water Resour. Res.* **2007**, *43*, W01420. [CrossRef]

111. Wang, X.; Xu, H.; Qian, B.; Feng, Y.; Mekis, E. Adjusted daily rainfall and snowfall data for Canada. *Atmos. Ocean.* **2017**, *55*, 155–168. [CrossRef]

112. Groisman, P.Y.; Koknaeva, V.V.; Belokrylova, T.A.; Karl, T.R. Overcoming biases of precipitation measurement: A history of the USSR experience. *Bull. Am. Meteorol. Soc.* **1991**, *72*, 1725–1732. [CrossRef]

113. Strangeways, I. Improving precipitation measurement. *Int. J. Climatol.* **2004**, *24*, 1443–1460. [CrossRef]

114. Adam, J.C.; Lettenmaier, D.P. Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res. Atmos.* **2003**, *108*, 4257. [CrossRef]

115. Kochendorfer, J.; Nitu, R.; Wolff, M.A.; Mekis, E.; Rasmussen, R.; Baker, B.; Earle, M.E.; Reverdin, A.; Wong, K.; Simith, C.D.; et al. Analysis of single-Alter-shielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3525–3542. [CrossRef]

116. Leeper, R.D.; Kochendorfer, J. Evaporation from weighing precipitation gauges: Impacts on automated gauge measurements and quality assurance methods. *Atmos. Meas. Tech.* **2015**, *8*, 2291–2300. [CrossRef]

117. Stewart, R.D.; Hut, R.; Rupp, D.; Gupta, H.; Selker, J.S. A resonating rainfall and evaporation recorder. *Water Resour. Res.* **2012**, *48*, W08601. [CrossRef]

118. Yang, D.; Ohata, T. A bias-corrected Siberian regional precipitation climatology. *J. Hydrometeorol.* **2001**, *2*, 122–139. [CrossRef]

119. Buishand, T. Some methods for testing the homogeneity of rainfall records. *J. Hydrol.* **1982**, *58*, 11–27. [CrossRef]

120. García-Marín, A.P.; Estévez, J.; Morbidelli, R.; Saltalippi, C.; Ayuso-Muñoz, J.L.; Flammini, A. Assessing inhomogeneities in extreme annual rainfall data series by multifractal approach. *Water* **2020**, *12*, 1030. [CrossRef]

121. Liu, W.; Fu, Z.; Chen, X.; Qu, J.; Wang, J.; Peng, X. Inhomogeneity of precipitation and its influencing factors in Northwest China from 1961 to 2015. *Theor. Appl. Climatol.* **2019**, *138*, 1831–1844. [CrossRef]

122. Peterson, T.C.; Easterling, D.R. Creation of homogeneous composite climatological reference series. *Int. J. Climatol.* **1994**, *14*, 671–680. [CrossRef]

123. Wu, Y.; Zhang, Z.; Huang, Y.; Jin, Q.; Chen, X.; Chang, J. Evaluation of the GPM IMERG v5 and TRMM 3B42 v7 Precipitation Products in the Yangtze River Basin, China. *Water* **2019**, *11*, 1459. [CrossRef]

124. Qiu, Q.; Liu, J.; Tian, J.; Jiao, Y.; Li, C.; Wang, W.; Yu, F. Evaluation of the radar QPE and rain gauge data merging methods in Northern China. *Remote Sens.* **2020**, *12*, 363. [CrossRef]

125. Thakur, M.K.; Kumar, T.V.L.; Rao, K.K.; Barbosa, H.; Rao, V.B. A new perspective in understanding rainfall from satellites over a complex topographic region of India. *Sci. Rep.* **2019**, *9*, 15610. [CrossRef] [PubMed]

126. Steiner, M.; Smith, J.A.; Burges, S.J.; Alonso, C.V.; Darden, R.W. Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.* **1999**, *35*, 2487–2503. [CrossRef]

127. 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]

128. Villarini, G.; Mandapaka, P.V.; Krajewski, W.F.; Moore, R.J. Rainfall and sampling uncertainties: A rain gauge perspective. *J. Geophys. Res. Atmos.* **2008**, *113*, D11102. [CrossRef]

129. Wood, S.J.; Jones, D.A.; Moore, R.J. Accuracy of rainfall measurement for scales of hydrological interest. *Hydrol. Earth Syst. Sci.* **2000**, *4*, 531–543. [CrossRef]

Microphysical and Polarimetric Radar Signatures of an Epic Flood Event in Southern China

Yu Ma [1], Haonan Chen [2,3], Guangheng Ni [1], V. Chandrasekar [2], Yabin Gou [4,*] and Wenjuan Zhang [5]

[1] Department of Hydraulic Engineering, State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, China; mayu15@mails.tsinghua.edu.cn (Y.M.); ghni@tsinghua.edu.cn (G.N.)

[2] Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA; haonan.chen@colostate.edu (H.C.); chandra@colostate.edu (V.C.)

[3] NOAA Physical Sciences Laboratory, Boulder, CO 80305, USA

[4] Hangzhou Meteorological Bureau, Hangzhou 310051, China

[5] State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China; zwj@cma.gov.cn

* Correspondence: 2017011201@cuit.edu.cn; Tel.: +86-13376812207

Abstract: An extremely heavy rainfall event hit Guangdong province, China, from 27 August to 1 September 2018. There were two different extreme rain regions, respectively, at the Pearl River estuary and eastern Guangdong, and a record-breaking daily precipitation of 1056.7 mm was observed at Gaotan station on 30 August. This paper utilizes a suite of observations from soundings, a gauge network, disdrometers, and polarimetric radars to gain insights to the two rainfall centers. The large-scale meteorological forcing, rainfall patterns, and microphysical processes, as well as radar-based precipitation signatures are investigated. It is concluded that a west-moving monsoon depression played a critical role in sustaining the moisture supply to the two extreme rain regions, and the combined orographic enhancement further contributed to the torrential rainfall over Gaotan station. The raindrop size distributions (DSD) observed at Zhuhai and Huidong stations, as well as the observed polarimetric radar signatures indicate that the rainfall at Doumen region was characterized by larger raindrops but a lower number concentration compared with that at Gaotan region. In addition, the dual-polarization radars are used to quantify precipitation intensity during this extreme event, providing timely information for flood warning and emergency management decision-making.

Keywords: extreme rainfall; polarimetric radar signatures; quantitative precipitation estimation; southern china

1. Introduction

Torrential rainfall events are one of the most severe disasters around the world [1,2]. The extreme rainfall and induced floods, landslides, debris flows gravely threaten life and property. The precipitation microphysics such as raindrop size distribution (DSD) serves as a fundamental bridge in deriving radar quantitative precipitation estimation (QPE) algorithms, which is critical for improving the accuracy of precipitation estimation and predictions [3,4]. Accurate precipitation estimates are also important input to the flash flood guidance systems for flood forecast, as well as subsequent warning operations and emergency management decision-making [5]. Therefore, a better understanding of precipitation microphysics and accurate quantitative precipitation estimation for extreme rain are important for flood warning and emergency management decision-making.

Under the influence of the East Asian monsoon, the Indian monsoon, the western Pacific subtropical high, as well as Tibetan Plateau, southern China is severely affected by heavy rain events during warm seasons (May to September), which usually cause floods and landslides [6–8]. From 27 August to 1 September 2018, an extremely heavy rainfall occurred over Guangdong province, especially in the south and east parts of Guangdong. There were two different extreme rainfall centers on 29 August and 30 August, respectively: one was located around Doumen station in Zhuhai, and the other was located at Gaotan station in Huidong. A record-breaking daily precipitation of 1056.7 mm was observed at Gaotan station on 30 August. This heavy rainfall caused catastrophic floods in many cities such as Huizhou, Shantou, Zhuhai, affecting more than 1 million people, causing directed financial losses around USD 144 million (https://www.thepaper.cn/newsDetail_forward_2404157). Moreover, a recent study has shown that extreme precipitation shows an increasing trend in south China during the last several decades [9], which highlights the importance of accurate precipitation measurement and modeling.

However, it is always a challenge to obtain accurate precipitation estimation. Gauges, weather radars, and satellite-based sensors are three main methods to measure precipitation [10]. Gauges can provide the most direct and precise precipitation observations. However, they are limited to fixed locations, and the networks of gauges are sparse. Therefore, interpolation is required to produce areal rainfall mapping and the interpolation method could lead to significant errors [11]. Satellite rainfall data has the advantage of large-scale spatial coverage, so the derived spatial distribution of rainfall is more complete. However, the satellite data also suffers from various sources of errors, including systematic error, random error, etc. Additionally, the spatial and temporal resolutions are very low, which has posed great difficulty in capturing the structure and evolution of small scale but strong storms [12]. In such cases, the only practical way to achieve a comprehensive estimation of precipitation is weather radar, which can provide real-time high-resolution monitoring over large areas [13,14].

Eight weather radars in the Guangdong area have completed dual-polarization upgrades in 2017 to improve disaster warning and forecasting capabilities. Compared to traditional single-polarization radar, the dual-polarization radar can measure polarimetric parameters including differential reflectivity Z_{dr}, differential phase shift ϕ_{dp}, and co-polar correlation coefficient ρ_{HV}[14]. These parameters can be used to reveal microphysical properties of different hydrometeors [15–18] and improve quantitative precipitation estimation [14,19]. Meanwhile, several disdrometers have been installed in Guangdong province. Though they provide point measurements, the accumulation of time decreases the spatial variability of local precipitation microphysics [20]. Therefore, these disdrometers can provide detailed knowledge of local DSD information, which is critical in understanding the microphysical characteristics of precipitation and improving microphysical parameterization schemes in the numerical weather prediction models [4,21].

The primary purpose of this study is to conduct a comprehensive analysis of this extreme event from 27 August to 1 September 2018, based on various in situ and remote sensing observations including rain gauges, polarimetric radars, disdrometers, and reanalysis data so as to gain a better understanding of the epic flood events as such, especially to explore the potential of polarimetric radars to resolve the microphysics and quantify the precipitation. This study is also part of our effort in improving precipitation monitoring, forecast, and associated hydrologic responses in southern China. The paper is organized as follows. The study domain and dataset are described in Section 2. The synoptic environment of this extreme rainfall event is detailed in Section 3. The rainfall pattern and microstructural characteristics of this rainstorm observed by gauge and disdrometers, as well as the associated polarimetric radar signatures are detailed in Section 4. Section 5 summarizes the main findings of this study and suggests future directions of this research.

2. Data and Methodology

2.1. Data

The observational data used in this study include rainfall measurements from a dense gauge network, two S-band polarimetric radars, and two second-generation Particle Size and Velocity (Parsivel^2) disdrometers in Guangdong. The instrument locations and the two special gauge stations (Gaotan and Doumen) are shown in Figure 1. The National Centers for Environmental Prediction (NCEP) final operational model global analysis data (NCEP-FNL) available every 6 h with a resolution of 0.25° × 0.25° at 31 vertical levels (http://rda.ucar.edu/datasets/ds083.2//#!access) are used to resolve the synoptic condition [22], along with the sounding data collected at 00:00 and 12:00 UTC at Qingyuan (QY) and Shantou (ST) (No. 59280, and No. 59316 from the University of Wyoming: http://weather.uwyo.edu/upperair/sounding.html).
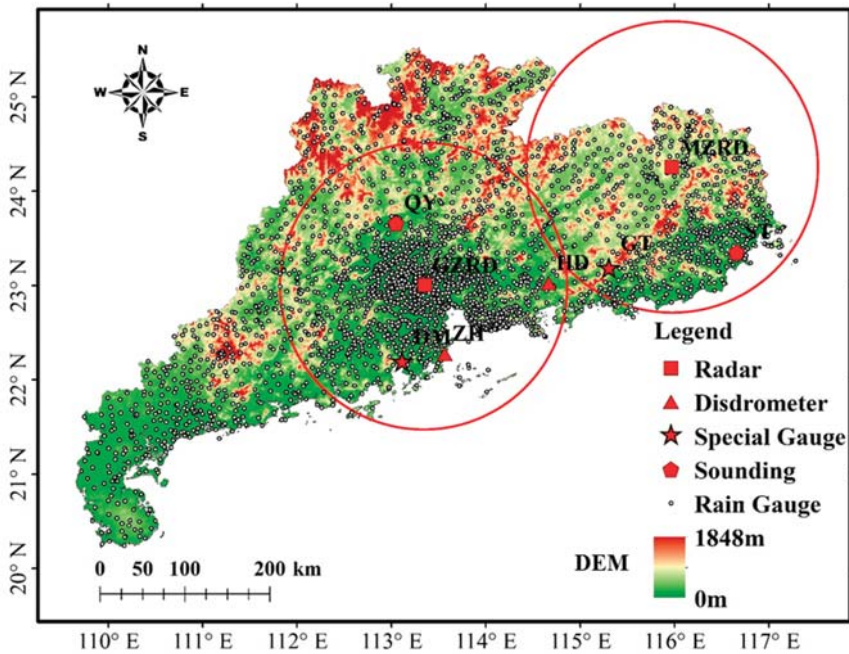
Figure 1. The topography of Guangdong province and geographical locations of instruments used in this study. The red circles are the 150-km coverage range rings from the radars. The red square, triangle, star, and pentagon represent the locations of radar, disdrometer, special gauge, and sounding, respectively. The small black circles stand for the gauges. The instrument names are abbreviated version of location names: DM (Doumen), GT (Gaotan), HD (Huidong), QY (Qingyang), ST (Shantou), ZH (Zhuhai).

2.2. Raindrop Size Distribution

In order to understand the microphysics of the extreme rain, two disdrometers nearest to Doumen and Gaotan gauge stations are used to provide the DSD observation. These two disdrometers are located at Zhuhai and Huidong, collocated with rain gauges within 20 m. Both disdrometers are optical disdrometers with a 54 cm^2 horizontal sample area and are configured with 1-min sampling resolution to measure the DSD and fall velocity of raindrops [23,24]. The disdrometer performance has been assessed and improved since it was invented, and many previous studies have been conducted with this device [25–28]. In particular, the velocity and particle sizes are divided into 32 non-uniform

bins, varying from 0.05 to 20.8 m s^{-1} for bin-center velocity and 0.062–24.5 mm for bin-center diameter (for detailed information in the user manual, https://www.manualslib.com/products/Ott-Parsivel2-5889584.html). The direct measurements from disdrometer are the number of raindrops at each velocity (*i*) and diameter (*j*) bin. Here, we take the bin-center value of each bin as the corresponding value. Several parameters used to describe the characteristics of DSD are calculated in the following.

The total number of raindrops can be calculated as follows:

$$Td = \sum_{i=1}^{32} \sum_{j=1}^{32} n_{i,j}, \tag{1}$$

where $n_{i,j}$ is the number of drops at each bin.

The number concentration of raindrops per unit volume for the *j*th diameter bin $N(D_j)$ can be calculated as follows:

$$N(D_j) = \sum_{i=1}^{32} \frac{n_{i,j}}{A \cdot \Delta t \cdot V_i \cdot \Delta D_j}, \tag{2}$$

where $N(D_j)$ is in m^{-3} mm^{-1}; A is the sampling area in m^2; Δt is the sampling time interval in s; A and Δt are, respectively, 0.0054 m^2 and 60 s in this study; ΔD_j (mm) is the diameter interval from D_j to D_{j+1} for the jth diameter bin; V_i (m s^{-1}) is the fall speed for the *i*th velocity class. Due to the measurement error, especially for larger size drops [23], the empirical terminal velocity–diameter ($V - D$) relation in Atlas et al. [29] is adopted in this study:

$$V(D_j) = 9.65 - 10.3 \exp\left(-0.6D_j\right), \tag{3}$$

The drops with velocity out the range of ±60% $V(D_j)$ are removed from the analysis [30].

The total number concentration N_t (m^{-3}), the mass weighted diameter D_m (mm), and normalized intercept parameter N_w (m^{-3} mm^{-1}) [14] are derived as:

$$N_t = \sum_{i=i}^{32} \sum_{j=1}^{32} \frac{n_{i,j}}{A \cdot \Delta t \cdot V_i}, \tag{4}$$

$$D_m = \frac{\sum_{j=1}^{32} N(D_j) \cdot D_j^4 \cdot \Delta D_j}{\sum_{j=1}^{32} N(D_j) \cdot D_j^3 \cdot \Delta D_j}, \tag{5}$$

$$N_w = \frac{4^4}{\pi \rho_w} \left(\frac{10^3 W}{D_m^4}\right), \tag{6}$$

D_m is closely related to the drop size; N_t and N_w are related to the number of raindrops. All these parameters are important in representing the DSD characteristics.

The integral rainfall parameters including rain rate R (mm h^{-1}) and liquid water content W (g m^{-3}) are calculated based on the following equations:

$$R = \frac{6\pi}{10^4 \rho_w} \sum_{j=1}^{32} V(D_j) D_j^3 N(D_j) \Delta D_j, \tag{7}$$

$$W = \frac{\pi \rho_w}{6 \times 10^3} \sum_{j=1}^{32} D_j^3 N(D_j) \Delta D_j, \tag{8}$$

where ρ_w is the water density (1.0 g cm^{-3}).

Meanwhile, a series of polarimetric radar variables are simulated at S-band frequency based on the DSD measurements using the T-matrix method [31–33], including horizontal reflectivity Z_h (mm^6 m^{-3}, or Z_H in dBZ), differential reflectivity Z_{dr} (dB), and specific differential phase K_{dp} (degree km^{-1}). The drop shape model used in the simulation is the one proposed by Thurai et al. [34] and temperature is 20 °C. The canting angle is not taken into account (i.e., canting angle is 0) as the DSD measurements are near ground.

We also want to note that to minimize the measurement errors and improve data reliability, several quality control procedures were applied on the 1-min DSD data. First, because of the low signal-to-noise ratios, the first two diameter bins are always empty, so the data in first two bins are eliminated in the analysis [23]. Second, the 1-min sample data with total raindrop number smaller than 10 or the derived rain rate less than 0.1 mm h^{-1} are considered noise and removed [23]. Then, if the continuous data satisfying the above conditions last less than 5 min, they will be ignored to avoid the spurious and erratic measurements [25,35]. Additionally, threshold on simulated radar parameter (i.e., Z_H < 55 dBZ) is used to further guarantee the creditability of the measured DSD data. The DSDs with the radar parameters out of the range are deleted to avoid mixed phase hydrometeors.

2.3. Radar Quantitative Precipitation Estimation

The two S-band radars are located at Guangzhou (hereafter referred as GZRD, 23.004° N, 113.355° E, 179 m) and Meizhou (MZRD, 24.256° N, 115.975° E, 423 m). Both radars are configured with a 6-min time resolution and 250-m range gate spacing, and have undergone rigorous quality control to ensure the data quality [36]. These two radars are used to monitor the evolution of the storm system and associated microphysical signatures. Moreover, these two radars are applied to estimate the rainfall to show the great potential of radar quantitative precipitation estimation.

For polarimetric radar, $R(Z_H)$, $R(K_{dp})$, $R(Z_H, Z_{dr})$, $R(Z_{dr}, K_{dp})$ are the four relations commonly used to estimate rainfall [14]. The parameters (i.e., coefficients and exponents) in these relations are determined by the local precipitation microphysics, and are usually derived from the in situ DSD data. The representative of these parameters and how to combine different relations are the key issue in deriving radar QPE [14].

In this study, four rainfall algorithms are applied to quantify the precipitation intensity and amounts during this event. These algorithms belong to two categories: one is $R(Z_H)$ relation, i.e., WRS-88D Z-R relationship [13] and localized Z-R relationship; another is a combination of the four rainfall relations (i.e., $R(Z_H)$, $R(K_{dp})$, $R(Z_H, Z_{dr})$, $R(Z_{dr}, K_{dp})$), i.e., the "adapted algorithm" described by Xia et al. [36] and the localized blended relation. Both Z-R relations are commonly used for single-polarized radar. The "adapted algorithm" is derived from DSD data and adjusted based on gauge observation and it has been demonstrated during several typhoon events in Southern China, which showed great performance [36].

Based on the nonlinear least-square method with the DSD data from two stations, the localized relations are fitted as follows:

$$R(Z_H) \ = \ 0.06082 \times 10^{0.05709Z_H}, \tag{9}$$

$$R(K_{dp}) \ = \ 40.4615 K_{dp}^{0.7703} \tag{10}$$

$$R(Z_H, Z_{dr}) \ = \ 0.00632 \times 10^{0.09134Z_H} 10^{-0.3325Z_{dr}} \tag{11}$$

$$R(Z_{dr}, K_{dp}) \ = \ 84.4318 K_{DP}^{0.9377} 10^{-0.1588Z_{dr}} \tag{12}$$

where Equation (9) is referred to as the localized Z-R relation, and the localized blended relation is a combination of Equations (10)–(12), using the same logic described by Xia et al. [36].

3. Synoptic Environment during This Epic Rainfall Event

From 27 to 31 August 2018, a monsoon depression evolved, moving west from the East China Sea. The resultant precipitation system was moving from west to east along the coast, causing heavy precipitation nearly over the whole Guangdong province, especially in the south and southeast parts of Guangdong on 29 and 30 August. The synoptic evolution of this precipitation system was analyzed using the NCEP-FNL data at 12:00 UTC 29 August and 12:00 UTC 30 August (Figure 2). Additionally, Figure 3 shows the stratification curve and convective available potential energy (CAPE) based on the sounding data at QY and ST stations.
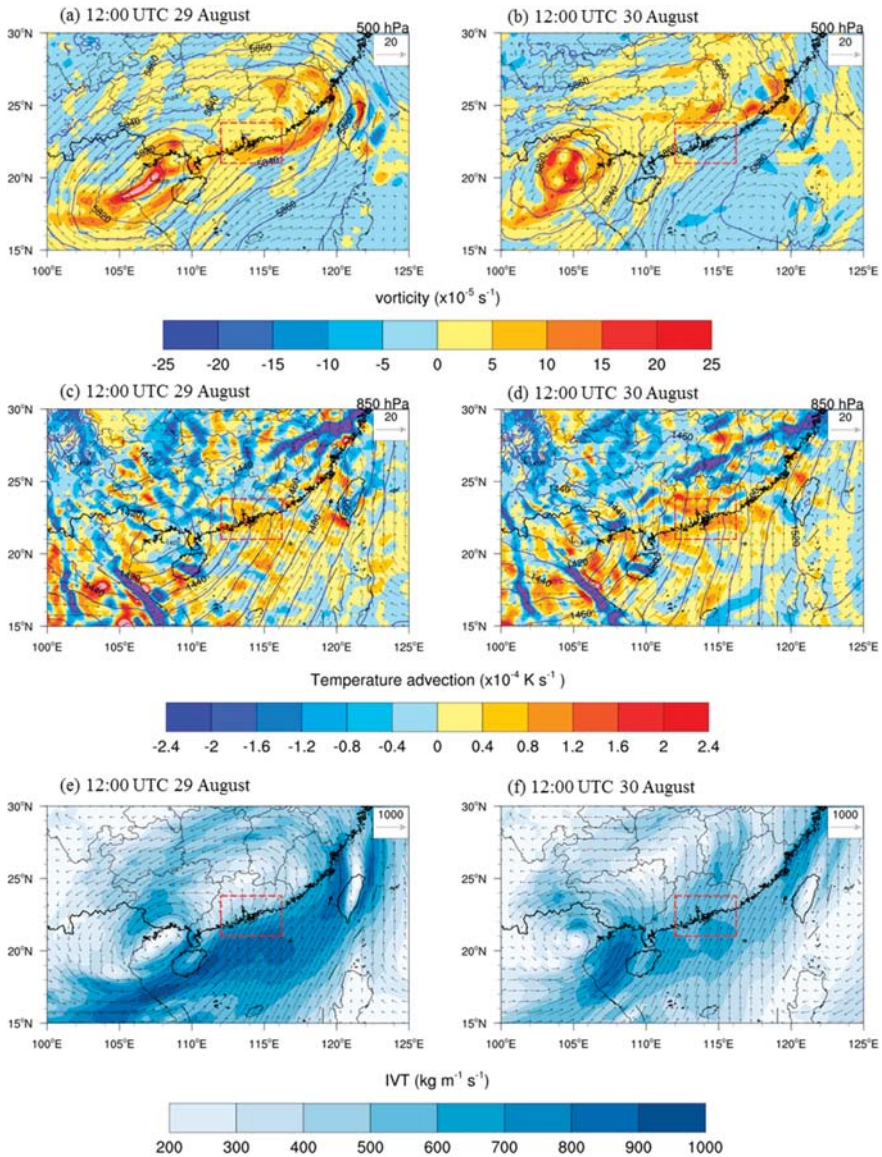
Figure 2. Geopotential height (gpm, solid blue lines at intervals of 10 gpm), winds (m s^{-1}, black vector arrows, vector scale at upper right), and vorticity (shaded, ×10^{-4} s^{-1}) at 500 hPa (**a**) at 12:00 UTC 29 August, and (**b**) at 12:00 UTC 30 August. The red (blue) shadings in panels (**a**,**b**) indicate positive (negative) vorticity. Geopotential height (gpm, solid blue lines at intervals of 10 gpm), winds (m s^{-1}, black vector arrows, vector scale at upper right), and the temperature advection (shaded, ×10^{-4} °C s^{-1}) at 850 hPa (**c**) at 12:00 UTC 29 August, and (**d**) at 12:00 UTC 30 August. The red (blue) shadings in panels (**c**) and (**d**) indicate warm (cold) temperature advection. The 1000–300 hPa vertically integrated water vapor transport (IVT, shaded, kg m^{-1}s^{-1}, also in black vector arrows, vector scale at upper right) (**e**) at 12:00 UTC 29 August, and (**f**) at 12:00 UTC 30 August. The dashed red rectangle indicates the area of the southeast of Guangdong province.
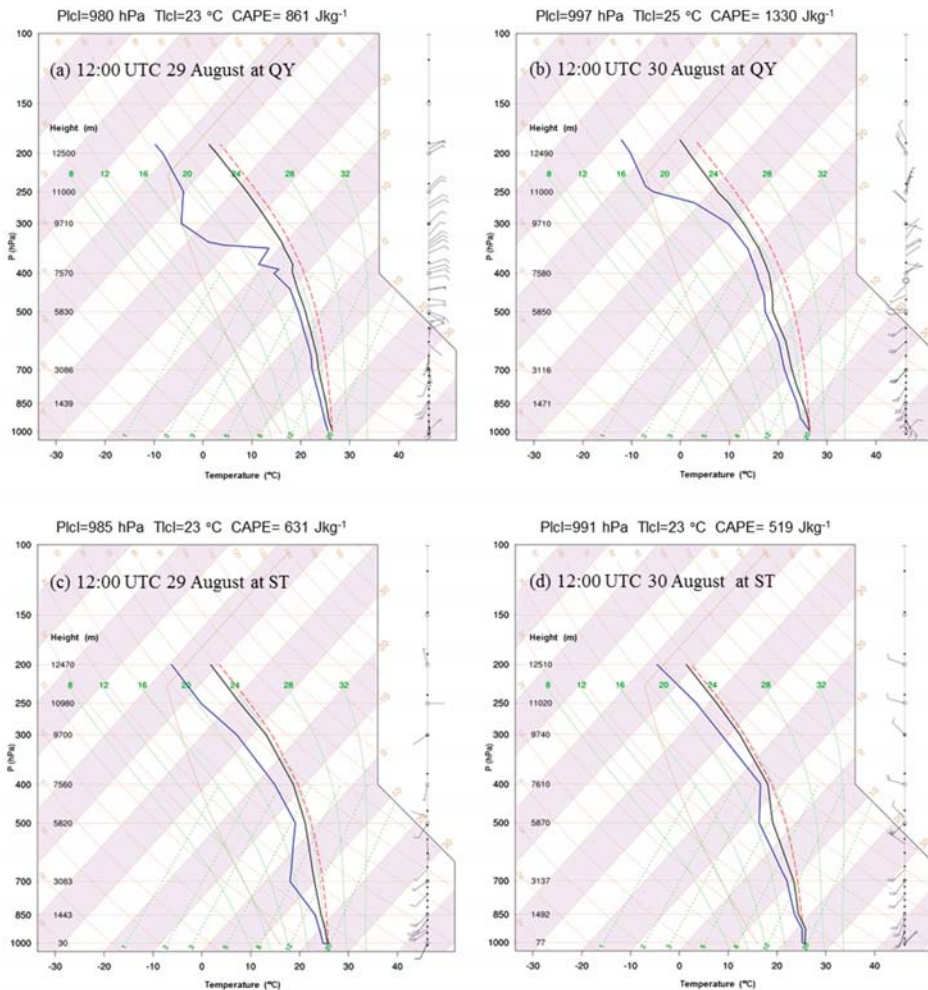
Figure 3. The skew T-log10P sounding profiles at (**a**,**b**) QY station and (**c**,**d**) ST station at 1200 UTC 29 August (left) and 1200 UTC 30 August (right). The red dashed line and black and blue solid lines represent the stratification curve, temperature profile, and dew point profile, respectively.

It can be seen that the entire region of southern China was under the influence of a monsoon depression across the mid- and lower-troposphere on 29 August. The depression moved west, bringing strong moistures transported to Guangdong province driven by the persistent south and southwest wind (Figure 2a,b). In the coastal areas of Guangdong, warm advection from the ground and lower troposphere continued to transport warm air to inland areas (Figure 2c,d). Due to the combined evolution of depression and topography lifting effect, the precipitation system was further strengthened on 30 August. The wind and temperature changes between Guangdong and the adjacent South China Sea were more violent (Figure 2c,d). Meanwhile, the vertically integrated vapor transport (IVT) over the coastal zone was much stronger on 30 August than that on 29 August (Figure 2e,f).

Figure 3a illustrates the sounding curves at 12:00 UTC 29 August at QY station. The large convective available potential energy (CAPE, ≈860 J kg^{-1}) and low level of free convection (LFC, ≈980 hPa) suggested that the atmospheric environment was favorable for the development of convection [37].

At 20:00 UTC 30 August, the CAPE became larger and LFC became lower, indicating that the convection had further developed. The soundings at ST station also showed a conducive environment for convection development though the CAPE decreased slightly from the 29th to 30th.

4. Precipitation Analysis Results and Discussion

4.1. Precipitation Pattern Observed by a Gauge Network

Figure 4 shows the spatial distribution of the cumulative rainfall in Guangdong during the period from 00:00 UTC 27 August to 00:00 UTC 2 September based on the gauge hourly precipitation data using Inverse Distance Weighted (IDW) interpolation method. This result was only used for visualization so as to get a general sense of the geographic distribution of rainfall. Obviously, this rainstorm affected most regions of Guangdong, with the extreme rainfall occurring at the southeast part and gradually decreased to the northwest of Guangdong. The gauge maximum cumulative rainfall was 1394.6 mm recorded at Gaotan station (23.1883 °N, 115.3044 °E), which set a new record of cumulative rainfall during a single event in Guangdong. Meanwhile, the maximum daily rainfall of 1056.7mm was also recorded at Gaotan station from 21:00 UTC 29 August to 21:00 UTC 30 August, which broke the historical daily rainfall record of 924.3 mm on 17 August 2013 [7].
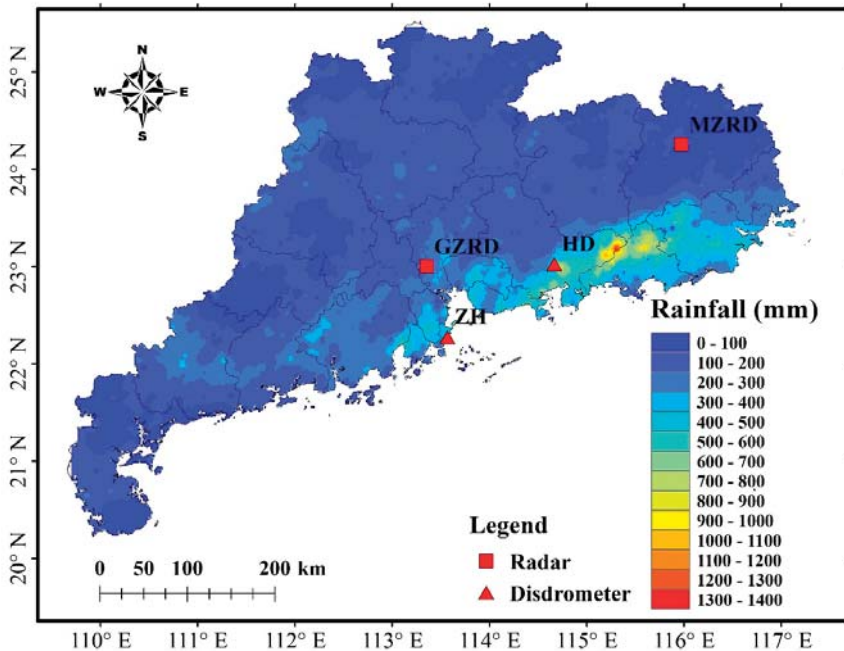
Figure 4. The spatial distribution of the accumulated rainfall from 27 August 2018 to 1 September 2018 of Guangdong province. The red squares and triangles represent the locations of radars and disdrometers, respectively.

Figure 5 shows the distribution of the daily rainfall that exceeded 100 mm from 28 to 31 August. It is clear that the rainfall is mainly concentrated on 29 and 30 August. There is a clear rain band along the coast on 29 August, with a gradient of accumulated rainfall decreasing from the coast towards the inland region, which can be partially attributed to the land-sea roughness contrasts [38]. On 30 August, the rain band moved a little towards northeast and inland region, with extremely high accumulated value (1041.1 mm) concentrated at Gaotan station, which could be due to the combined effect of terrain

and evolved synoptic condition. To further reveal the two rainfall processes, Figure 6 shows the time series of rainfall collected at two gauges where extreme daily rainfall concentrated on these two days. One station is located at Doumen (22.1967 °N, 1153.1131 °E) in Zhuhai, representing the rainfall on 29 August, when the daily rainfall reached 458.6 mm. Another station is Gaotan station, representing the most extreme rainfall pattern both for 30 August and the whole event.
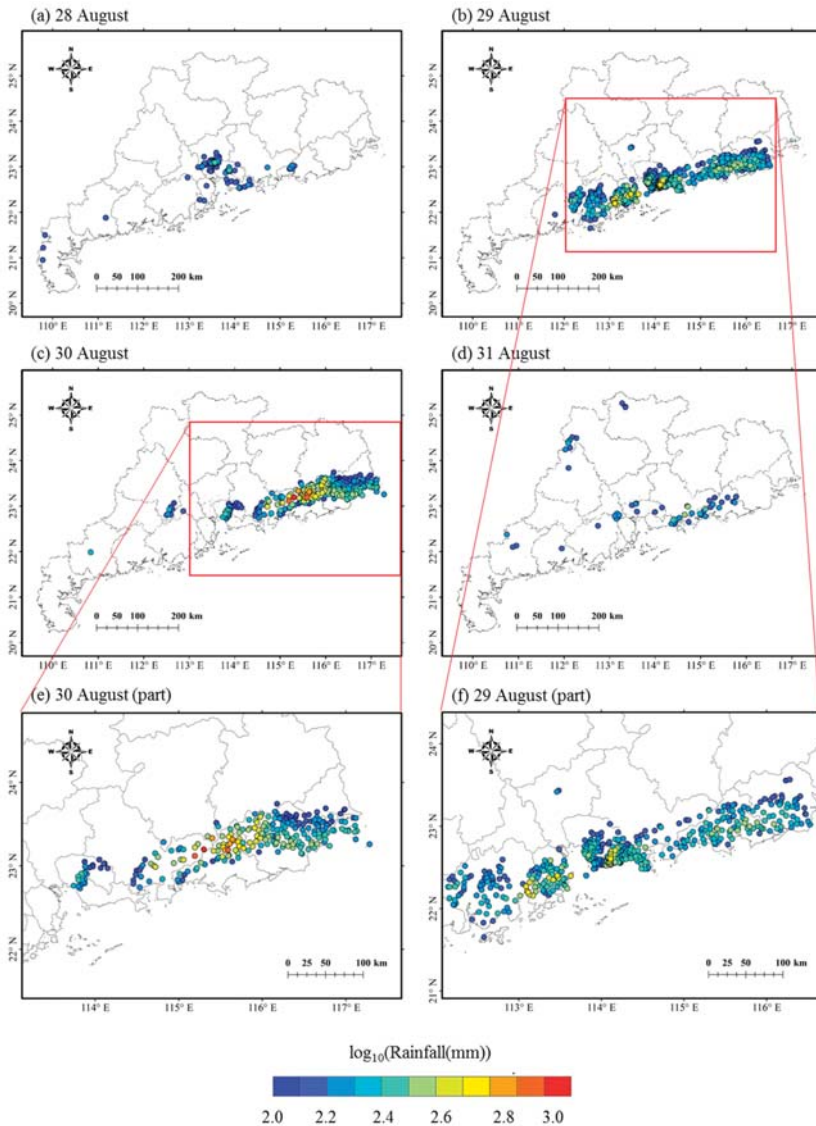
Figure 5. Daily rainfall accumulation (in mm, a day is defined from 0000 UTC to 0000 UTC of the next day) for (**a**) 28 August, (**b**) 29 August, (**c**) 30 August, (**d**) 31 August, (**e**) a zoomed area for 30 August, (**f**) a zoomed area for 29 August. Each dot represents a gauge station with daily rainfall accumulation exceeding 100 mm.
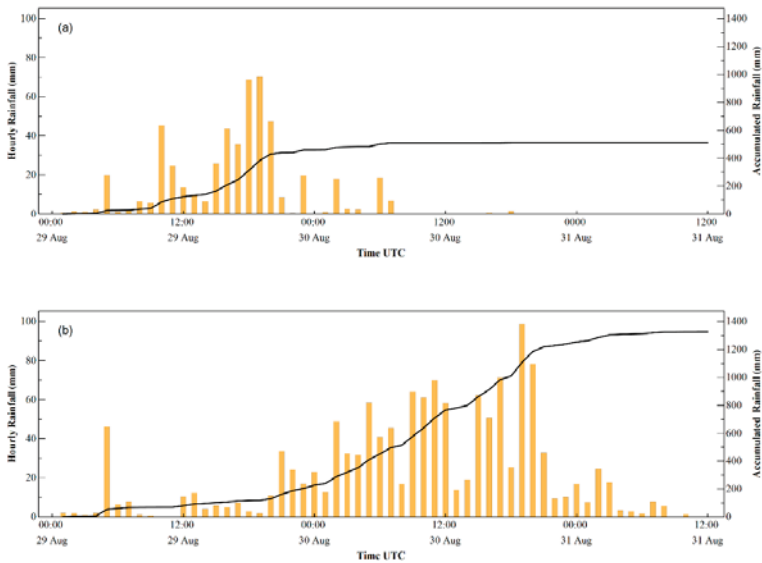
Figure 6. Hourly and accumulated precipitation at (**a**) Doumen station, (**b**) Gaotan station from 00:00 UTC 29 August to 12:00 UTC 31 August.

The differences between the two stations are clear: Doumen station experienced heavy rainfall mainly on 29 August, while Gaotan station experienced a heavier and longer rainfall from late 29 August to early 31 August. Doumen station experienced three major rainfall stages on 29 August: first from 00:00 to 06:00 UTC, second from 06:00 to 14:00 UTC, and third from 14:00 to 23:00 UTC. The third stage was longer and stronger than the first two stages, with a maximum hourly rainfall of 70.2 mm from 1800 to 1900 UTC 29 August. After 0800 UTC 30 August, rainfall at Doumen station almost disappeared. At Gaotan station, rainfall exhibited two major stages: first from 00:00 to 20:00 UTC on 29 August, then from 20:00 UTC 29 August to 08:00 UTC 31 August. The first stage, with only one hourly rainfall exceeding 20 mm is much weaker than the second. During the second stage, the hourly rainfall showed an increasing trend, and reached a maximum of 98.6 mm from 1900 to 2000 UTC 30 August. Moreover, the cumulative rainfall of the second stage (1206.6 mm) accounted for 87% rainfall of the whole event at Gaotan station.

4.2. Raindrop Size Distribution

4.2.1. DSDs Time Series at Two Observation Stations

As there are no disdrometers deployed at the Doumen and Gaotan stations, two nearby Parsivel[2] disdrometers were selected to analyze the raindrop size distribution from 00:00 UTC 29 August to 12:00 UTC 31 August. One disdrometer is located at Zhuhai (22.2750 °N, 113.5669 °E) near the Doumen station (45 km away); the other is located at Huidong (23.0261 °N, 114.6681 °E) near the Gaotan station (67 km away) (see Figure 1). The comparisons of hourly rainfall between the disdrometers and the gauges at Zhuhai and Huidong stations (within 20 m for each pair) are shown in Figure 7. Although the disdrometer might slightly underestimate the rainfall compared with gauge data because of the absence of some records and the overlap of drops along the laser beam, overall, they captured the rainfall pattern very well. As such, we ignored the impact on DSD observations in this study. Moreover, the rainfall pattern at Zhuhai station is similar to that of Doumen station, and the rainfall pattern of Huidong station is similar to Gaotan station. The similar patterns indicate that these two disdrometers could, respectively, represent the DSD characteristic of Doumen and Gaotan stations.
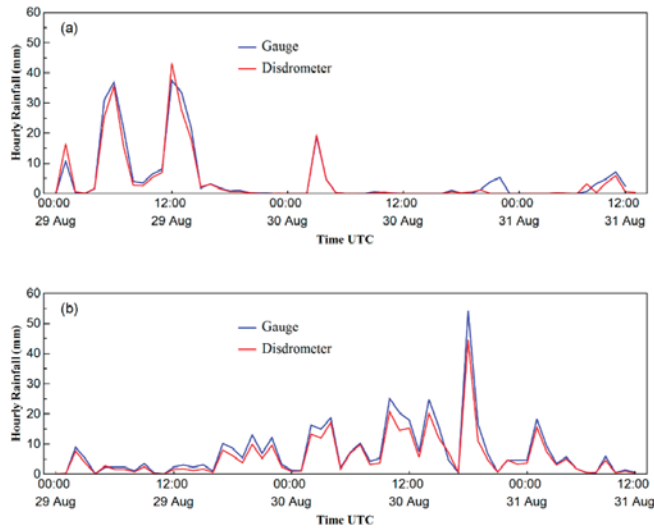
Figure 7. Hourly rainfall measured by gauges and collocated disdrometers: (**a**) Zhuhai station, (**b**) Huidong station from 00:00 UTC 29 August to 12:00 UTC 31 August.

Figure 8 shows the time series of the drop size distribution ($N(D)$: m^{-3} mm^{-1}) in the logarithmic scale from 00:00 UTC 29 August to 12:00 UTC 31 August, derived by the 1-min disdrometer observations at the two stations. In line with the synoptic condition shown in Figure 2, it can be seen that the Zhuhai station experienced stronger rainfall during 29 August, and it reached the maximum hourly rainfall of about 40 mm (Figure 7). The maximum raindrops greater than 7 mm were recorded from 12:00 to 14:00 UTC 29 August. After that, rainfall at Zhuhai station began to weaken with the decreasing number of raindrop concentration and size of raindrops, while that at Huidong station gradually increased. After 23:00 UTC 29 August, this long-lasting storm disappeared at Zhuhai station following with two short storms characterized with lower concentration but bigger raindrops around 04:00 UTC 30 August and around 11:00 UTC 31 August. The rain at Huidong station showed an increasing trend from 13:00 UTC 29 August, and the number of raindrop concentration and size of raindrops also show an increasing trend, and the hourly rainfall reached the highest from 18:00 to 19:00 UTC 30 August. Compared to the extreme rainfall period at Zhuhai station, the number of large drops ($D > 5$ mm) at Huidong station is much less during the strong rainfall period which we hypothesize is due to the collision caused by sea-inland orographic effect [38–40].

4.2.2. The Distribution of D_m and N_w

Figure 9a shows the scatterplot of $\log_{10} N_w$ vs. D_m of the two stations. The maritime convective clusters, continental convective clusters, and the stratiform rain line described in Bringi et al. [41] as well as the convection-stratiform separation line suggested by Thompson et al. [42] are also shown in Figure 9a. The scatter points of two stations show similar boundaries, of which there are both convective rain and stratiform rain types. Meanwhile, there are more points classified as stratiform rain by the classification suggested by Thompson et al. [42], and the averaged $\log_{10} N_w$-D_m pairs of two stations are close to the stratiform line described by Bringi et al. [41]. Moreover, for convective rain type, there are more points located at the continental convective cluster rather than maritime cluster for both stations. Compared with the DSD characteristics studied at Yangjiang from July to August by Tang et al. [43], the $\log_{10} N_w$ of both two stations are smaller no matter what rain types, suggesting a lower drop concentration of this rain system.
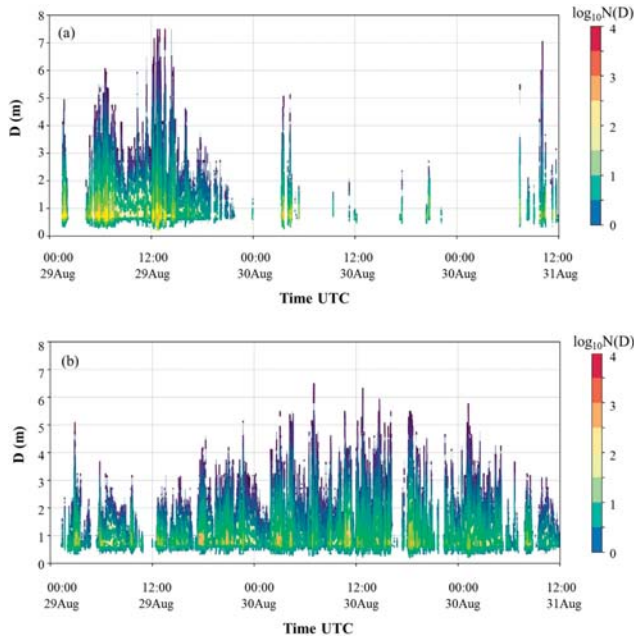
Figure 8. Time series of raindrop size distributions (DSD) observed at (**a**) Zhuhai station, (**b**) Huidong station from 00:00 UTC 29 August to 12:00 UTC 31 August. The color density in both (**a**) and (**b**) represents the number concentration in logarithmic units of $N(D)(\mathrm{m}^{-3}\,\mathrm{mm}^{-1})$.
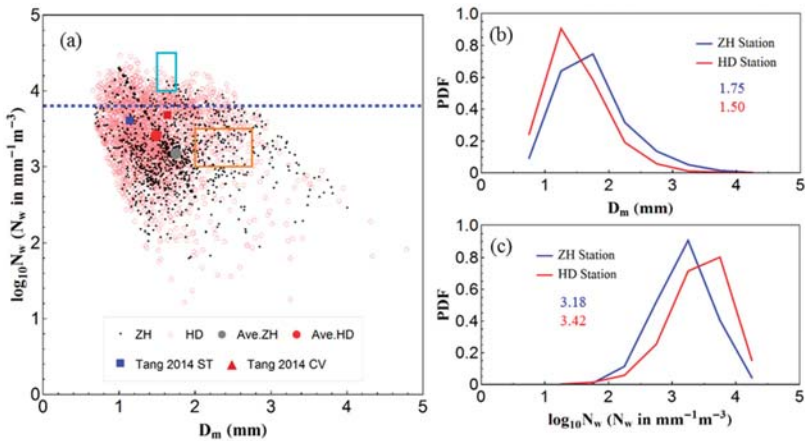
Figure 9. (**a**) The scatterplots of $\log_{10} N_w$ vs. D_m; (**b**) probability density functions (PDFs) of D_m; (**c**) PDFs of $\log_{10} N_w$ at Huidong and Zhuhai stations. Both data are from 00:00 UTC 29 August to 12:00 UTC 31 August. The black dot and pink circle represent Zhuhai data and Huidong data, respectively. The mean values of $\log_{10} N_w$ and D_m are represented by circles (gray for Zhuhai and red for Huidong). The two rectangles correspond to the maritime (cyan) and continental convective (orange) clusters reported by Bringi et al. [41]. The black and purple dash-dot lines represent the characteristics of stratiform rain described in Bringi et al. [41] and the convection/stratiform separation line from Thompson et al. [42], respectively. The square symbols represent the average values of various types of rain, from Tang et al. [43]. CV stands for convective rain, ST is stratiform rain. ZH is short for Zhuhai, and HD is short for Huidong. It is the same in the following figures and tables.

The probability density functions (PDFs) of D_m and $\log_{10} N_w$ with the mean values are presented in Figure 9b,c. The PDF of D_m at Zhuhai station peaks around 1.25 mm while that at Huidong peaks around at 1.75 mm. Additionally, the PDF of $\log_{10} N_w$ at Zhuhai station peaks around 3.25 while that at Huidong peaks around 3.75. This leads to a higher mean value of D_m (1.75 mm) and lower mean value of $\log_{10} N_w$ (3.18) at Zhuhai station than those at Huidong station. This result indicates that the rain at Zhuhai station has larger drops, but a lower drop number concentration compared to Huidong station.

4.2.3. The DSD Spectra

The characteristics of DSD of this rain system are investigated to understand the precipitation microphysical processes. To eliminate to the effect caused by rain rate and duration, the DSD measurements are divided into six classes according to the associated rain rate (R): C1, $0.1 \leq R < 1$; C2, $1 \leq R < 5$; C3, $5 \leq R < 10$; C4, $10 \leq R < 25$; C5, $25 \leq R < 50$ mm h^{-1}. Such classification is based on the rain rate distribution of this system, as well as several previous studies [25,35]. The DSD sample relative frequency, rain rate statistics and mean DSD parameters for each class are summarized in Tables 1 and 2. For lower rain rate classes (C1–C3), the rain rate of Huidong station is lower than that of Zhuhai station. For C5, the rain rates of two stations are close to each other and for the other two high rain rate classes (C4 and C6), the rain rate of Huidong station is higher than that of Zhuhai station, which caused the higher hourly rainfall at from 18:00 to 19:00 UTC 30 August. However, due to the lower relative frequency of high rain rate classes, the overall average rain rate of Huidong station is smaller than that of Zhuhai station.

Table 1. Number and DSD retrieved rain rate statistics of each rain rate class for Zhuhai and Huidong station data.

Classes	Rain Rate Threshold (mm h^{-1})	Relative Frequency		Mean (mm h^{-1})		SD (mm h^{-1})		Skewness	
		ZH	HD	ZH	HD	ZH	HD	ZH	HD
C1	$0.1 \leq R < 1$	0.32	0.32	0.52	0.48	0.25	0.26	0.05	0.27
C2	$1 \leq R < 5$	0.35	0.37	2.66	2.30	1.10	1.04	0.28	0.82
C3	$5 \leq R < 10$	0.09	0.11	7.41	7.07	1.45	1.44	0.17	0.32
C4	$10 \leq R < 25$	0.12	0.11	15.72	16.46	4.09	4.16	0.31	0.32
C5	$25 \leq R < 50$	0.09	0.06	35.41	35.38	6.85	7.30	0.25	0.34
C6	$R \geq 50$	0.03	0.02	68.14	71.02	12.14	16.40	0.77	0.87
All data	-	1	1	9.01	7.43	15.28	13.81	2.70	3.45

Table 2. Mean values DSD parameters for each rain rate class for Zhuhai and Huidong station data.

Classes	D_m (mm)		N_t (m^{-3})		$\log_{10} N_w$ (m^{-3} mm^{-1})		W (g m^{-3})	
	ZH	HD	ZH	HD	ZH	HD	ZH	HD
C1	1.29	1.14	54.8	81.0	2.94	3.18	0.030	0.031
C2	1.66	1.43	139.6	200.5	3.19	3.42	0.132	0.127
C3	1.96	1.64	255.2	409.6	3.30	3.67	0.337	0.360
C4	2.18	1.96	429.8	548.5	3.44	3.66	0.682	0.748
C5	2.62	2.32	648.2	786.8	3.44	3.66	1.409	1.474
C6	2.80	2.48	1051.5	1369.3	3.57	3.82	2.616	2.863
All data	1.75	1.50	232.1	286.8	3.18	3.42	0.378	0.337

It can be shown from Figure 10 that the total averaged DSD of Zhuhai station has wider spread as well as lower concentration at small and medium-sized drops ($D < 2$ mm) but higher concentration at big and large drops ($D > 2$ mm) compared with that of Huidong station, which corresponds to a larger D_m, more W, but a lower N_t for Zhuhai station compared with those for Huidong station (Table 2). For the same rain class, the DSD of Huidong shows the same characteristics with higher concentration at smaller drops as well as lower concentration at larger drops. Additionally, the diameters where the

raindrop concentrations of two stations become larger as the rain classes increased, from 1.2 to 3.8 mm, due to the higher the concentration of larger drops at higher rain rates for both stations.
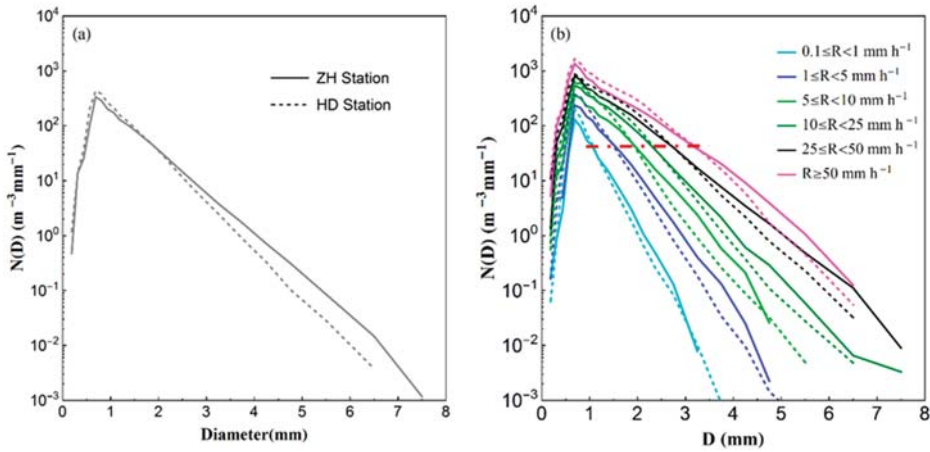
Figure 10. Mean DSDs observed at Zhuhai and Huidong stations from 00:00 UTC 29 August to 12:00 UTC 31 August: (**a**) all DSD data, (**b**) DSD data classified by different rain rates. The solid and dashed lines in (**a**) and (**b**) represent Zhuhai and Huidong stations, respectively. The dash-dot line in (**b**) shows the DSD intersection of two stations for each rain rate class.

Table 3 summarizes the relative contributions of each size class to N_t, W, and R, which are obtained through dividing the parameters calculated for the raindrops of each size class by the corresponding parameters calculated for all the size classes. For both stations, the small raindrops ($D < 1$ mm) have the highest concentration (51.99% and 57.92%) of the total data set and the second highest relative frequency (Td 38.08% and 44.26%), but relatively low contributions to total water content and rainfall. The contribution of small and medium-sized raindrops (D: 1–2 mm) to the total number is the greatest, accounting for 47.48% (45.99%) at Zhuhai (Huidong) station. However, for Zhuhai station, the middle-sized drops (D: 2–3 mm) accounts for the most rainfall (34.39%), while the small and medium-sized raindrops (D: 1–2 mm) account for the most for Huidong station (40.50%). Moreover, for Zhuhai station, the medium-sized and larger drops ($D > 2$ mm) account more for rainfall and water content compared with Huidong station, as the higher relative frequency. This result is consistent with the conclusion shown in Figures 8 and 9 that larger drops are less likely to occur at Huidong, where complex topography could exacerbate the collision.

Table 3. Relative contributions of each size class to the total drop numbers T_d, total drop concentration N_t, liquid water content W, rain rate R. The bold number highlights the maximum.

D (mm)	T_d (%)		N_t (%)		R (%)		W (%)	
	ZH	HD	ZH	HD	ZH	HD	ZH	HD
<1	38.08	44.26	**51.99**	**57.92**	3.91	6.30	7.79	11.69
1~2	**47.48**	**45.99**	39.54	36.60	30.70	**40.50**	**37.01**	**45.65**
2~3	11.67	8.41	7.03	4.82	**34.39**	34.00	31.33	28.83
3~4	2.24	1.16	1.18	0.58	19.22	13.80	15.25	10.16
>4	0.53	0.18	0.26	0.08	11.78	5.40	8.62	3.67

4.3. Polarimetric Radar Signatures and Rainfall Analysis

As the rain patterns of Zhuhai station and Huidong station are similar to that of Doumen station and Gaotan station, the DSD characteristics of Zhuhai and Huidong areas can partly represent the

microphysics of Doumen and Gaotan station, which suggests a high probability of a higher drop concentration at Gaotan station and a larger drops at Doumen station. However, the total rainfall difference between Huidong and Gaotan station is significant (≈400 mm vs. ≈1000 mm), which means the DSD of Gaotan may be different and need more observation to supplement the current study. In the following study, the dual-polarized radar observations are used to further understand the rainfall.

4.3.1. The Polarimetric Radar Signatures

To further understand the differences between the two rain patterns on 29 and 30 August, the polarimetric radar fields over the two extreme stations are compared. Figure 11 shows the time series of polarimetric variables over Doumen and Gaotan stations observed by GZRD and MZRD 0.5-degree sweeps, respectively. The size of radar pixel over Doumen (Gaotan) station is around 0.25×1.49 km (0.25×2.23 km), and the horizontal/vertical distances are 93 km/0.81 km (137 km/1.21 km) between the radar pixel and the station. Generally speaking, Figure 11 shows that the timing changes of radar variables are consistent with the rainfall changes at both stations (Figure 6). During the extreme rainfall periods of Doumen station and Gaotan station, the mean values of Z_H over both stations are similar while the mean value of Z_{dr} over Doumen station is higher than that over Gaotan station, indicating that bigger drops are more likely to occur at Doumen station than Gaotan station, which is similar to the DSD analysis of Zhuhai and Huidong.
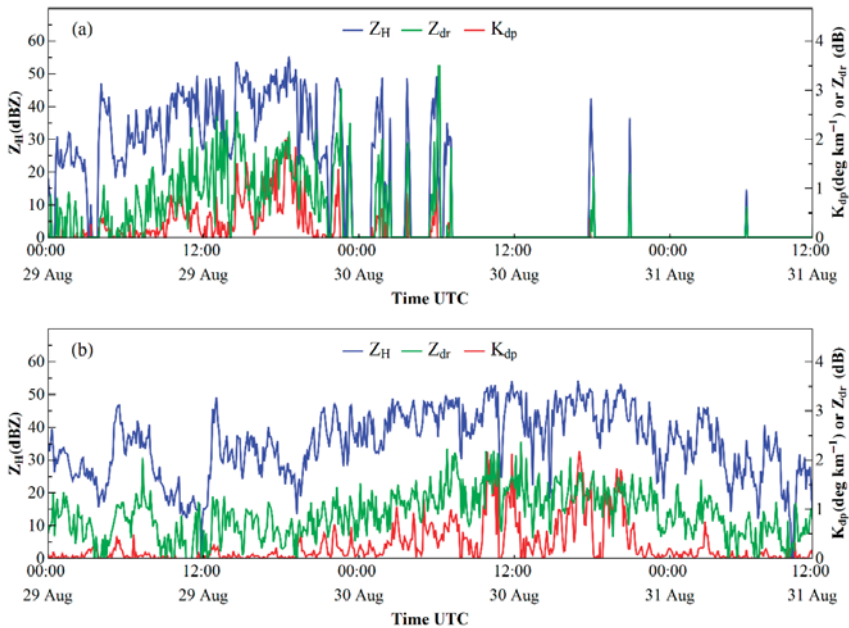
Figure 11. Time series of different polarimetric radar variables over (**a**) Doumen station observed by GZRD, (**b**) Gaotan station observed by MZRD from 00:00 UTC 29 August to 12:00 UTC 31 August.

To eliminate the random error of one radar pixel, two extreme regions were defined as circular areas with a radius of 5 km centered on Doumen station and Gaotan station, respectively, hereafter referred to Doumen rain region and Gaotan rain region. The polarimetric variables from 0.5-degree sweep over these two regions from 00:00 UTC 29 August to 12:00 UTC 31 August were analyzed. Figure 12 shows the joint frequency distributions of Z_{dr} vs. Z_H, K_{dp} vs. Z_H, as well as $\log_{10}(K_{dp}/Z_h)$ vs. Z_{dr} observed by radars over two regions as well as the simulated radar moments based on DSD data at Zhuhai station and Huidong station. The frequency at each pair of data bin (e.g., the bin width

is 1 dBZ for Z_H and 0.1 dB for Z_{dr}, 0.1 degree km^{-1} for K_{dp}, and 0.1 for $\log_{10}(K_{dp}/Z_h)$) refers to the total number of radar pixels with values of Z_{dr} vs. Z_H (K_{dp} vs. Z_H or $\log_{10}(K_{dp}/Z_h)$ vs. Z_{dr}) falling into the bins, and the frequency distribution is further normalized by dividing by the maximum frequency among all the pairs of data bins.
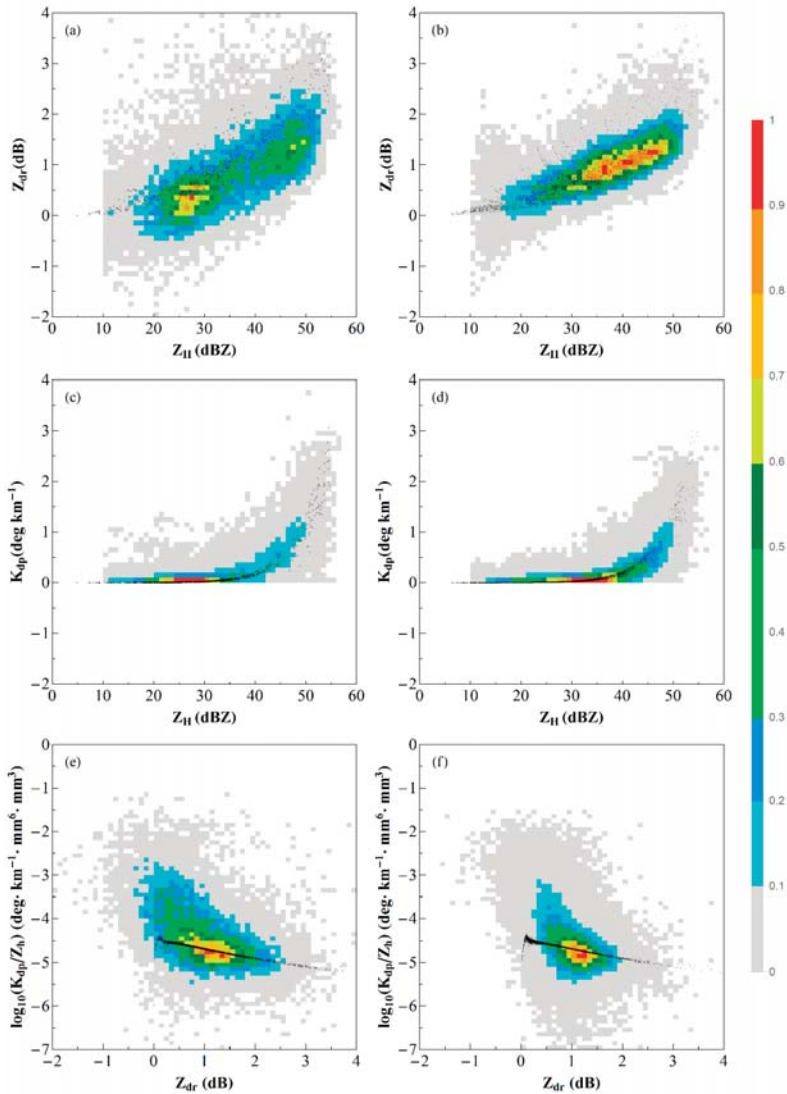
Figure 12. Joint frequency distributions of polarimetric radar at (**a**,**c**,**e**) Doumen and (**b**,**d**,**f**) Gaotan areas from 00:00 UTC 29 August to 00:00 UTC 31 August: (**a**,**b**) Z_{dr} vs. Z_H, (**c**,**d**) K_{dp} vs. Z_H, (**e**,**f**) $\log_{10}(K_{dp}/Z_h)$ vs. Z_{dr}. The black dots are the simulated radar moments based on DSD data at (**a**,**c**,**e**) Zhuhai station, (**b**,**d**,**f**) Huidong station.

Frequencies for the observed Z_H, Z_{dr} at Doumen rain region peaked at approximately 28 dBZ, and 0.3 dB, respectively. While at Gaotan rain region they peaked at high values, around 35 dBZ for Z_H and 0.8 dB for Z_{dr}. The Z_H maximum values with the 10% normalized frequency of two regions

are similar around 52 dBZ, which are consistent with the extreme rainfall periods of two stations. Meanwhile, the Z_{dr} maximum values with the 10% normalized frequency at Gaotan rain region is smaller (2.1 dB vs. 2.5 dB). Meanwhile, for the same Z_H, the range and mean Z_{dr} at Doumen rain region is larger than that at Gaotan station, which indicates that larger drops more likely occurred at Doumen compared with Gaotan. For both regions, the slope between Z_H and K_{dp} shows a sharp transition over sampling grids with Z_H exceeding 40 dBZ, resulting in a rapid increase of K_{dp} along with the increase of Z_H. Although the peak frequencies for K_{dp} at two rain regions are similar around 0.1 degree km^{-1}, K_{dp} increases faster as the Z_H increases at Doumen rain region which indicates that there are more relative larger drops at Doumen rain region. Then, considering the higher rain rate and smaller raindrops at Gaotan station, the number concentration must be larger. The extreme lower value of $\log_{10}(K_{dp}/Z_h)$ could confirm this, which means the concentration of small drops at Gaotan is higher than that at Doumen. It can be concluded from the polarimetric radar signatures that compared with Doumen rain region, the raindrops at Gaotan rain region have higher concentration but smaller drops, which is consistent with the DSD characteristic at Huidong station.

4.3.2. Radar-Based Quantitative Precipitation Estimation (QPE)

In this part, a set of various radar rainfall algorithms including the "adapted algorithm" [36], the localized blended algorithm, the localized relation and WSR-88D Z-R relation [13], were applied to quantify the precipitation intensity and amounts during this event. After estimating the instantaneous rainfall rates using various radar rainfall algorithms, the rainfall accumulations were computed at hourly scale. To quantify the performances of different algorithms, a set of metrics was computed, including the bias (*BIAS*), normalized mean bias (*NMB*), normalized mean absolute error (*NMAE*), Pearson's correlation coefficients (*CC*), which are defined as follows:

$$BIAS = \frac{\sum_{N=1}^{M}(R_N - G_N)}{M}, \tag{13}$$

$$NMB = \frac{[\sum_{N=1}^{M}(R_N - G_N)]/M}{(\sum_{N=1}^{M} G_N)/M} \times 100\%, \tag{14}$$

$$NMAE = \frac{(\sum_{N=1}^{M}|R_N - G_N|)/M}{(\sum_{N=1}^{M} G_N)/M} \times 100\%, \tag{15}$$

$$CC = \frac{\sum_{N=1}^{M}(R_N - \overline{R_N})(G_N - \overline{G_N})}{\sqrt{\sum_{N=1}^{M}(R_N - \overline{R_N})^2 \sum_{N=1}^{M}(G_N - \overline{G_N})^2}}, \tag{16}$$

where R_N and G_N represent the radar estimates of different algorithms and the rain gauge measurements at time frame N, respectively. M is the total sample number.

The comparison of estimates from different rainfall algorithms and rain gauge measurements at Gaotan and Doumen station is shown in Figure 13 and the evaluation results are shown in Table 4. As shown in Figure 13, all the algorithms have similar patterns to the rain gauge measurements, while QPE results of the "adapted algorithm" is better than other algorithms with lowest *BIAS*, *NMB*, *NMAE* and the highest *CC* (Table 4). This is in line with the findings during typhoon case studies [36]. Nevertheless, we should note that all the algorithms are underestimating the rain rates and accumulations during this flood event and the results still need to be optimized.
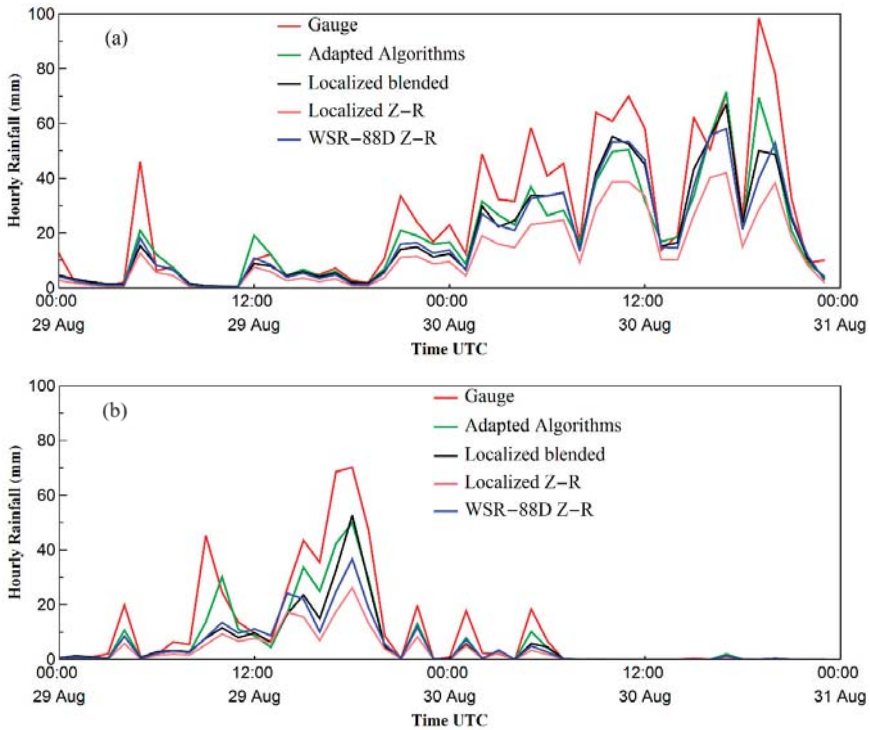
Figure 13. The rain gauge measurements (red line) and estimates computed by various radar algorithms. (**a**) Gaotan station, (**b**) Doumen station.

Table 4. Evaluation results of the various radar algorithms at Gaotan station and Doumen station.

Station	Metrics	Algorithm			
		"Adapted Algorithm"	Localized Blended Rainfall Algorithm	Localized Z–R Relation	WSR-88D Z-R Relation
Gaotan	BIAS (mm)	−6.61	−7.30	−13.17	−7.90
	NMB (%)	−25.37	−28.01	−50.56	−30.31
	NMAE (%)	29.64	30.16	50.56	32.26
	CC	0.95	0.94	0.93	0.93
Huidong	BIAS (mm)	−3.65	−4.92	−7.03	−5.47
	NMB (%)	−34.39	−46.37	−66.29	−51.63
	NMAE (%)	37.65	47.83	66.72	54.36
	CC	0.96	0.93	0.91	0.91

To further analyze the QPE results of various radar rainfall algorithms, the scatter plots of radar rainfall estimates versus gauge measurements at all rain gauge stations less than 100 km from the radars are shown in Figure 14 and the evaluation results are shown in Table 5. Most of the hourly rainfall rates are from 0 to 40 mm hr^{-1}. Again, the adapted algorithm has the best performance, while all the rainfall algorithms show underestimation compared to the gauge measurements, especially during heavy rain periods.
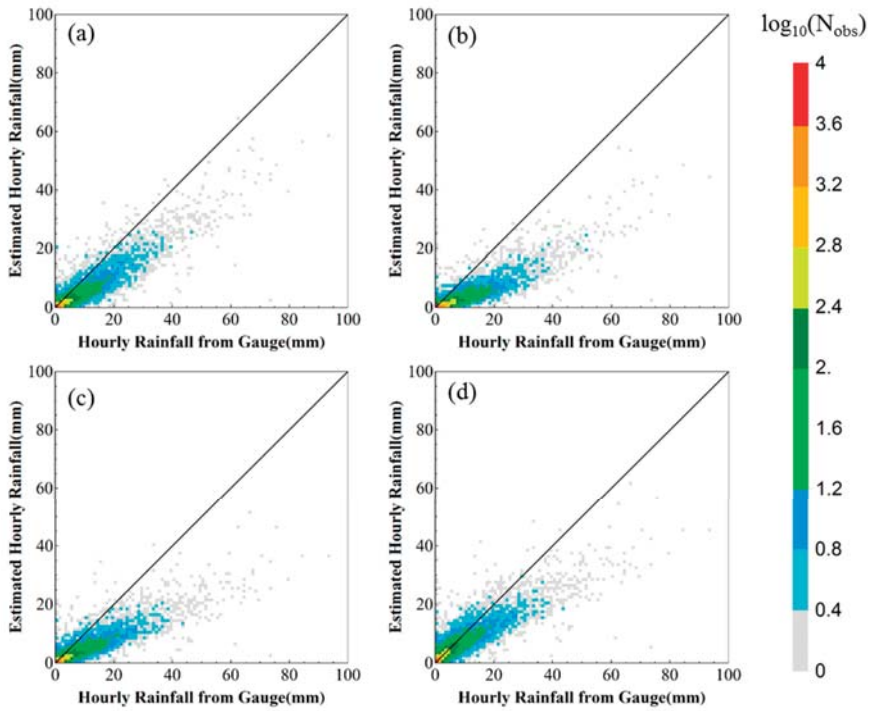
Figure 14. The scatter plots of radar estimated rainfall versus gauge measurements at gauge locations less than 100 km from the radar: (**a**) the "adapted algorithm" [36]; (**b**) localized blended rainfall algorithm; (**c**) localized Z–R relation; (**d**) WSR-88D Z-R relation [13]. The color density represents the observation sample numbers in logarithmic unit.

Table 5. Evaluation results of the various radar algorithms compared with gauges within 100-km scale of radars.

Metrics	Algorithm			
	"Adapted Algorithm"	Localized Blended Rainfall Algorithm	Localized Z–R Relation	WSR-88D Z-R Relation
BIAS (mm)	−0.29	−0.90	−1.38	−1.09
NMB (%)	−12.47	−39.21	−60.35	−47.70
NMAE (%)	39.93	46.38	63.24	53.76
CC	0.91	0.92	0.90	0.90

5. Discussion

Although the analysis of DSD and polarimetric radar signatures shows similar results, the relatively long distances between disdrometers and the extreme rain centers might induce some uncertainty in the representation of extreme rainfall DSD characteristics. As mentioned, the polarimetric radar signatures at the two rainfall centers are quite different, even for the same precipitation system, which is likely due to the complex falling processes as a result of orographic enhancement and dynamic cloud microphysics involved in this extreme event [44]. More disdrometers and in situ measurements would be required to fully resolve the three-dimensional structure of precipitation in such complex terrains.

In addition, although the polarimetric radars could provide more insights into the two extreme rain regions, the precipitation estimated by the current radar algorithms underestimate the rainfall compared to the ground gauges. The differences of sample areas between radars and gauges may be

one reason, especially when the gauges are far from the radars. In such cases, a network of short-range X-band polarimetric radars would be useful for better QPE. Additionally, the adapted algorithms have relatively better performance mainly because they are derived using DSD and gauge data collected during many storm events in Southern China [36]. However, at the same time, the DSD observed at both Huidong and Zhuhai stations during this event showed a lower number concentration compared with previous long-term studies in Yangjiang [43]. This reminds us that the rainfall algorithms should be appropriately developed based on local rainfall characteristics, which is still under investigation.

6. Conclusions

An epic flood event occurred in southern China from 27 August to 1 September 2018, with a maximum accumulative rainfall of 1394.6 mm recorded at Gaotan station. Multi-observations including reanalysis data, soundings, gauges, disdrometers, and polarimetric radars were used to comprehensively analyze this extreme precipitation event. Based on the reanalysis and gauge data, the synoptic environment and precipitation pattern were investigated. The DSD time series, the mass weighted diameter D_m, and normalized intercept parameter N_w of Huidong and Zhuhai stations were derived to achieve a better understanding of the precipitation microphysics. It was found that the depression coupled with land–sea interaction and orographic enhancement kept providing moisture to Guangdong province, leading to the extreme rainfall on 29 at Doumen station and on 30 August at Gaotan station. Although both locations are under the same synoptic system, the DSD observed at Huidong station showed a higher number concentration at smaller drop size and lower number concentration at larger drop size compared with the Zhuhai station. This is likely due to the stronger collision caused by sea–inland orographic effect near the Huidong station, which is similar to the microphysical signatures observed by polarimetric radars at Gaotan and Doumen rain regions. Future study is needed to reveal the mechanism of the sea–inland orographic effect on the precipitation microphysics. In addition, the radar based QPE results show that the rainfall algorithms have great influence on the accuracy. Extra work is required to improve local radar rainfall estimates in order to further help with the severe weather warning operations.

Author Contributions: Y.M. carried out the detailed analysis. H.C. and G.N. supervised the work. Y.M. and H.C. drafted the manuscript. G.N., V.C., and Y.G. reviewed and edited the manuscript. W.Z. collected the data and provided comments on this study. All authors have read and agreed to the published version of the manuscript.

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

References

1. Stevenson, S.N.; Schumacher, R.S. A 10-Year Survey of Extreme Rainfall Events in the Central and Eastern United States Using Gridded Multisensor Precipitation Analyses. *Mon. Weather Rev.* **2014**, *142*, 3147–3162. [CrossRef]
2. Zhou, T.; Yu, R.; Chen, H.; Dai, A.; Pan, Y. Summer Precipitation Frequency, Intensity, and Diurnal Cycle over China: A Comparison of Satellite Data with Rain Gauge Observations. *J. Clim.* **2008**, *21*, 3997–4010. [CrossRef]
3. Battan, L.J. *Radar Observation of the Atmosphere*; University of Chicago Press: Chicago, IL, USA, 1973.
4. Abel, S.J.; Boutle, I.A. An improved representation of the raindrop size distribution for single-moment microphysics schemes. *Quart. J. R. Meteor. Soc.* **2012**, *138*, 2151–2162. [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.* **2017**, *14*, 1208–1212. [CrossRef]

6. Huang, Y.; Liu, Y.; Liu, Y.; Knievel, J.C. Budget Analyses of a Record-Breaking Rainfall Event in the Coastal Metropolitan City of Guangzhou, China. *J. Geophys. Res. Atmos.* **2019**, *124*, 9391–9406. [CrossRef]

7. Li, H.; Wan, Q.; Peng, D.; Liu, X.; Xiao, H. Multiscale analysis of a record-breaking heavy rainfall event in Guangdong, China. *Atmos. Res.* **2020**, *232*, 104703. [CrossRef]

8. Chen, X.; Zhao, K.; Xue, M. Spatial and temporal characteristics of warm season convection over Pearl River Delta region, China, based on 3 years of operational radar data. *J. Geophys. Res. Atmos.* **2014**, *119*, 412–447. [CrossRef]

9. Fu, S.; Li, D.; Sun, J.; Si, D.; Ling, J.; Tian, F. A 31-year trend of the hourly precipitation over South China and the underlying mechanisms. *Atmos. Sci. Lett.* **2016**, *17*, 216–222. [CrossRef]

10. Li, Z.; Yang, D.; Hong, Y. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *J. Hydrol.* **2013**, *500*, 157–169. [CrossRef]

11. Ma, Y.; Zhang, Y.; Yang, D.; Farhan, S.B. Precipitation bias variability versus various gauges under different climatic conditions over the Third Pole Environment (TPE) region. *Int. J. Climatol.* **2015**, *35*, 1201–1211. [CrossRef]

12. Wang, Z.; Zhong, R.; Lai, C.; Chen, J. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.* **2017**, *196*, 151–163. [CrossRef]

13. Fulton, R.A.; Breidenbach, J.P.; Seo, D.; Miller, D.A.; O'Bannon, T. The WSR-88D Rainfall Algorithm. *Weather Forecast* **1998**, *13*, 377–395. [CrossRef]

14. Bringi, V.N.; Chandrasekar, V. *Polarimetric Doppler Weather Radar: Principles and Applications*; Cambridge University Press: Cambridge, UK, 2001.

15. Bechini, R.; Chandrasekar, V. A Semisupervised Robust Hydrometeor Classification Method for Dual-Polarization Radar Applications. *J. Atmos. Ocean. Tech.* **2015**, *32*, 22–47. [CrossRef]

16. Park, H.S.; Ryzhkov, A.V.; Zrnić, D.S.; Kim, K. The Hydrometeor Classification Algorithm for the Polarimetric WSR-88D: Description and Application to an MCS. *Weather Forecast* **2009**, *24*, 730–748. [CrossRef]

17. Wen, G.; Chen, H.; Zhang, G.; Sun, J. An Inverse Model for Raindrop Size Distribution Retrieval with Polarimetric Variables. *Remote Sens.* **2018**, *10*, 1179. [CrossRef]

18. Sun, Y.; Xiao, H.; Yang, H.; Feng, L.; Chen, H.; Luo, L. An Inverse Mapping Table Method for Raindrop Size Distribution Parameters Retrieval Using X-band Dual-Polarization Radar Observations. *IEEE Trans. Geosci. Remote Sens.* **2020**, 1–22. [CrossRef]

19. Ryzhkov, A.V.; Giangrande, S.E.; Schuur, T.J. Rainfall Estimation with a Polarimetric Prototype of WSR-88D. *J. Appl. Meteorol.* **2005**, *44*, 502–515. [CrossRef]

20. Berne, A.; Jaffrain, J.; Schleiss, M. Scaling analysis of the variability of the rain drop size distribution at small scale. *Adv. Water Resour.* **2012**, *45*, 2–12. [CrossRef]

21. McFarquhar, G.M.; Hsieh, T.; Freer, M.; Mascio, J.; Jewett, B.F. The Characterization of Ice Hydrometeor Gamma Size Distributions as Volumes in N_0–λ–μ Phase Space: Implications for Microphysical Process Modeling. *J. Atmos. Sci.* **2015**, *72*, 892–909. [CrossRef]

22. Kalnay, E. The NCEP/NCAR Reanalysis 40-year Project. B. *Am. Meteorol. Soc.* **1996**, *77*, 437–471. [CrossRef]

23. Tokay, A.; Wolff, D.B.; Petersen, W.A. Evaluation of the New Version of the Laser-Optical Disdrometer, OTT Parsivel2. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1276–1288. [CrossRef]

24. Löffler-Mang, M.; Joss, J. An Optical Disdrometer for Measuring Size and Velocity of Hydrometeors. *J. Atmos. Ocean. Tech.* **2000**, *17*, 130–139. [CrossRef]

25. Ma, Y.; Ni, G.; Chandra, C.V.; Tian, F.; Chen, H. Statistical characteristics of raindrop size distribution during rainy seasons in the Beijing urban area and implications for radar rainfall estimation. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 4153–4170. [CrossRef]

26. Luo, L.; Xiao, H.; Yang, H.; Chen, H.; Guo, J.; Sun, Y.; Feng, L. Raindrop size distribution and microphysical characteristics of a great rainstorm in 2016 in Beijing, China. *Atmos. Res.* **2020**, *239*, 104895. [CrossRef]

27. Janapati, J.; Seela, B.K.; Reddy, M.V.; Reddy, K.K.; Lin, P.; Rao, T.N.; Liu, C. A study on raindrop size distribution variability in before and after landfall precipitations of tropical cyclones observed over southern India. *J. Atmos. Sol. Terr. Phys.* **2017**, *159*, 23–40. [CrossRef]

28. Seo, B.; Krajewski, W.F.; Quintero, F.; El Saadani, M.; Goska, R.; Cunha, L.K.; Dolan, B.; Wolff, D.B.; Smith, J.A.; Rutledge, S.A.; et al. Comprehensive Evaluation of the IFloodS Radar Rainfall Products for Hydrologic Applications. *J. Hydrometeor.* **2018**, *19*, 1793–1813. [CrossRef]

29. Atlas, D.; Srivastava, R.C.; Sekhon, R.S. Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.* **1973**, *11*, 1–35. [CrossRef]

30. Jaffrain, J.; Berne, A. Experimental Quantification of the Sampling Uncertainty Associated with Measurements from PARSIVEL Disdrometers. *J. Hydrometeorol.* **2010**, *12*, 352–370. [CrossRef]

31. Leinonen, J. High-level interface to T-matrix scattering calculations: Architecture, capabilities and limitations. *Opt. Express* **2014**, *22*, 1655. [CrossRef]

32. Mackowski, D.W.; Mishchenko, M.I. A multiple sphere T-matrix Fortran code for use on parallel computer clusters. *J. Quant. Spectrosc. Radiat. Transf.* **2011**, *112*, 2182–2189. [CrossRef]

33. Waterman, P.C. Matrix formulation of electromagnetic scattering. *Proc. IEEE* **1965**, *53*, 805–812. [CrossRef]

34. Thurai, M.; Huang, G.J.; Bringi, V.N.; Randeu, W.L.; Schönhuber, M. Drop Shapes, Model Comparisons, and Calculations of Polarimetric Radar Parameters in Rain. *J. Atmos. Ocean. Technol.* **2007**, *24*, 1019–1032. [CrossRef]

35. Jash, D.; Resmi, E.A.; Unnikrishnan, C.K.; Sumesh, R.K.; Sreekanth, T.S.; Sukumar, N.; Ramachandran, K.K. Variation in rain drop size distribution and rain integral parameters during southwest monsoon over a tropical station: An inter-comparison of disdrometer and Micro Rain Radar. *Atmos. Res.* **2019**, *217*, 24–36. [CrossRef]

36. Xia, Q.; Zhang, W.; Chen, H.; Lee, W.-C.; Han, L.; Ma, Y.; Liu, X. Quantification of Precipitation Using Polarimetric Radar Measurements during Several Typhoon Events in Southern China. *Remote Sens.* **2020**, *12*, 2058. [CrossRef]

37. Weisman, M.L.; Klemp, J.B.; Rotunno, R. Structure and Evolution of Numerically Simulated Squall Lines. *J. Atmos. Sci.* **1988**, *45*, 1990–2013. [CrossRef]

38. Yang, L.; Smith, J.; Liu, M.; Baeck, M.L. Extreme rainfall from Hurricane Harvey (2017): Empirical intercomparisons of WRF simulations and polarimetric radar fields. *Atmos. Res.* **2019**, *223*, 114–131. [CrossRef]

39. Powell, M.D. Changes in the low-level kinematic and thermodynamic structure of hurricane Alicia (1983) at landfall. *Mon. Weather Rev.* **1987**, *115*, 75–99. [CrossRef]

40. Cifelli, R.; Chandrasekar, V.; Chen, H.; Johnson, L.E. High resolution radar quantitative precipitation estimation in the San Francisco Bay Area: Rainfall monitoring for the urban environment. *J. Meteorol. Soc. Jpn. Ser. II* **2018**, *96A*, 141–155. [CrossRef]

41. Bringi, V.N.; Chandrasekar, V.; Hubbert, J.; Gorgucci, E.; Randeu, W.L.; Schoenhuber, M. Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *J. Atmos. Sci.* **2003**, *60*, 354–365. [CrossRef]

42. Thompson, E.J.; Rutledge, S.A.; Dolan, B.; Thurai, M. Drop size distributions and radar observations of convective and stratiform rain over the equatorial Indian and West Pacific Oceans. *J. Atmos. Sci.* **2015**, *72*, 4091–4125. [CrossRef]

43. Tang, Q.; Xiao, H.; Guo, C.; Feng, L. Characteristics of the raindrop size distributions and their retrieved polarimetric radar parameters in northern and southern China. *Atmos. Res.* **2014**, *135–136*, 59–75. [CrossRef]

44. Chen, H.; Cifelli, R.; White, A. Improving operational radar rainfall estimates using profiler observations over complex terrain in Northern California. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 1821–1832. [CrossRef]

Regional Precipitation Model Based on Geographically and Temporally Weighted Regression Kriging

Wei Zhang [1,*], Dan Liu [1], Shengjie Zheng [2], Shuya Liu [1], Hugo A. Loáiciga [3] and Wenkai Li [1]

[1] School of Geography and Information Engineering, China University of Geosciences (Wuhan),
    Wuhan 430074, China; cugld@cug.edu.cn (D.L.); cug_lsy@cug.edu.cn (S.L.); 20151003928@cug.edu.cn (W.L.)
[2] China Petroleum Pipeline Engineering CO., LTD., Langfang 065000, China; sj_zheng@cug.edu.cn
[3] Department of Geography, University of California, Santa Barbara, CA 93106, USA; hloaiciga@ucsb.edu
* Correspondence: weizhang@cug.edu.cn

Abstract: High-resolution precipitation field has been widely used in hydrological and meteorological modeling. This paper establishes the spatial and temporal distribution model of precipitation in Hubei Province from 2006 through 2014, based on the data of 75 meteorological stations. This paper applies a geographically and temporally weighted regression kriging (GTWRK) model to precipitation and assesses the effects of timescales and a time-weighted function on precipitation interpolation. This work's results indicate that: (1) the optimal timescale of the geographically and temporally weighted regression (GTWR) precipitation model is daily. The fitting accuracy is improved when the timescale is converted from months and years to days. The average mean absolute error (MAE), mean relative error (MRE), and the root mean square error (RMSE) decrease with scaling from monthly to daily time steps by 36%, 56%, and 35%, respectively, and the same statistical indexes decrease by 13%, 15%, and 14%, respectively, when scaling from annual to daily steps; (2) the time weight function based on an exponential function improves the predictive skill of the GTWR model by 3% when compared to geographically weighted regression (GWR) using a monthly time step; and (3) the GTWRK has the highest accuracy, and improves the MAE, MRE and RMSE by 3%, 10% and 1% with respect to monthly precipitation predictions, respectively, and by 3%, 10% and 5% concerning annual precipitation predictions, respectively, compared with the GWR results.

Keywords: precipitation interpolation; geographically and temporally weighted regression; time weight function; geographically and temporally weighted regression kriging

1. Introduction

Precipitation is a critical flux in the water cycle [1,2]. It is, for this reason, imperative to study the spatial–temporal features of precipitation [3,4]. Precipitation data are usually derived from meteorological sites with limited spatial coverage and sensor-gathered data, such as remote sensing satellites and rainfall radars [5]. Meteorological site location observations yield local, discrete, and limited spatial data points, which cannot account for the spatial precipitation variability accurately [6,7]. The general spatial resolution of remote sensing precipitation data products is generally low, which does capture the precipitation distribution in small areas [8]. These problems constrain the application of precipitation data for multiple practical purposes. Thus, there is a need for further study on how to obtain continuous and accurate distributions of precipitation at regional scales.

Spatial interpolation of precipitation data falls into two categories: deterministic interpolation and spatial–temporal interpolation. Deterministic interpolation is further divided into two categories: global interpolation and local interpolation [9–11]. Spatial–temporal interpolation includes two categories:

subtraction and extension [12]. Spatial–temporal interpolation methods for spatial–temporal irregular dataset interpolation and missing data patching include the spatial–temporal inverse distance weighting method, the spatial–temporal kriging method, and collaborative spatial-temporal kriging, among the main ones [13–15]. It has become a common practice to explore the distribution of precipitation employing spatial statistical analysis to cope with the spatial-temporal non-smoothness of precipitation. The geographically weighted regression (GWR) model was proposed for the study of spatial relations and spatial correlation, based on the common linear regression model by Fotheringham et al. [16]. The GWR model prescribes parameter estimation based on the location function expressing the non-stationary spatial features of precipitation. The regression coefficients in the GWR model capture the locational attribute. They can, therefore, take into account the influence of spatial heterogeneity, thus significantly improving the ability to analyze the variation in spatiotemporal characteristics of precipitation. This means the GWR model has attracted wide attention regarding quantitative precipitation estimation, as well as other spatial variables [17,18]. However, the GWR model only considers the spatial characteristics of precipitation data, while ignoring time characteristics of precipitation. The geographically and temporally weighted regression (GTWR) model was proposed in 2010 by Huang et al., and incorporates the time dimension into the model formulation [19]. On the one hand, the GTWR model has the basic characteristics of a general variable coefficient model and exhibits the high fitting skill of the local regression model, which captures the differences in spatial position and takes into account the spatial heterogeneity of precipitation. On the other hand, the model adds the time series traits, synthesizes the time dimension distribution information of the sample points, and embeds the spatial–temporal characteristics into the model [20].

GTWR performs wells in predicting spatial–temporal heterogeneity, and many studies in a variety of fields of science have proven the effectiveness of the GTWR model in spatial economic analysis, atmospheric sciences, population analysis, and other social and economic fields. The GTWR model was applied to model housing price data in London by Fotheringham et al. [21], which validated the proposed method and its superiority over the traditional GWR method while highlighting the importance of time explicit spatial modeling. The GTWR model was applied to assess the spatial–temporal differences in the influence of each driving factor on the scale of carbon emissions and the intensity of carbon emissions in China by Xiao et al. [22]. Liu et al. studied housing price data and related factors in Beijing from 1980 to 2016 [23], to propose a calculation method for travel distance, applying the GWR. The GTWR model was employed to study the influencing factors on housing prices, and it was concluded that the GTWR model is suitable for identifying effective real estate management policies. The fire record data from 2002 to 2010 in Hefei, China, was reviewed by Song et al. [24], using the linear model (LM), GWR, and GTWR to model urban fire risk. The latter authors concluded that road density and commercial spatial distribution have the most significant influence on fire risk. GTWR can detect small changes in variable spatial–temporal heterogeneity of diverse phenomena. The performance of the GTWR model was verified with particulate matter ≤2.5 μm (PM2.5) concentration data in the Xuzhou area, China, and compared with ordinary least squares (OLS), GWR, and time-weighted regression (TWR) models by Bai et al. [25]. The results indicate that the regression coefficient of the GTWR model was the highest, and its interpolation skill was optimal. The GTWR model was applied to estimate the ground concentration of nitrogen dioxide (NO_2) in central China by Qin et al. [26], and cross-validation results proved that the fitting results of the GTWR model were better than those of the OLS, GWR, and TWR models. Five models, including GTWR, were implemented to analyze the relationships between PM2.5 and other criteria of air pollutants by Wei et al. [27], and GTWR showed great advantages over the other three models in terms of higher model R^2 and more desirable model residuals, and only slightly less than TWR.

Precipitation has a high causal correlation both in space and time. Therefore, it is intuitively logical to use the GTWR model to fit precipitation data. At present, the use of spatial statistical analysis to fit precipitation interpolation is mainly represented by the GWR model. Brunsdon et al. [28] reported a study of the relation between total annual precipitation and elevation in the UK by employing the

GWR model. Their results revealed that the rate of precipitation increased with elevation, and that the predicted sea level precipitation varied between 600 mm and 1250 mm. The precipitation data from the Tropical Rainfall Measuring Mission (TRMM) 3B43 products were fitted with a multi-variable GWR reduction method to obtain 1 km × 1 km precipitation data by Chen et al. [29]. The GWR method was compared with two other downscaling methods (single variable regression (UR) and multivariate regression (MR)). Chen et al. (29) concluded that the GWR method could predict annual and monthly TRMM 1 km x 1 km precipitation with high precision. The accuracy of TRMM precipitation products at the daily and monthly scales in the Qaidam Basin of China was evaluated by Lv et al. [30] with the GWR model. Their results indicate that the precipitation GWR model based on ground and satellite data reduced the error of TRMM products, which was of significance in the fields of hydrology and climate change. The vegetation and climate data (Normalized Vegetation Difference Index (NVDI) and rainfall) from 2002 through 2012 for the growing season (June–September) in the Sahel region of Africa was relied upon in the GWR model by Georganos et al. [31]. The results showed that the spatial pattern of the NDVI–rainfall relationship is characterized when selecting the appropriate scale. Their GWR model performs better than the OLS in terms of predictive skill, accuracy, and residual autocorrelations. With the further research of scholars, geographically weighted regression kriging (GWRK) [32], as an extension of the GWR model, appears in the spatial interpolation of temperature and soil properties. It has also been explored in the field of precipitation and prediction research, achieving excellent results. The GWRK model combines the GWR with the kriging method, and uses the kriging method to interpolate the residual part of the GWR model, which eliminates the influence of the spatial correlation of the residual on the model fit, and shows that it is masked by spatial non-stationary local variation.

The GTWR model focuses primarily on the time dimension, although it accounts for the characterization of spatial heterogeneity [33]. At a particular timescale, the GTWR handles the distribution of the time dimension in a manner dissimilar to that described by the first law of geography (Tobler [34]), and, thus, it can be improved. Ge et al. proposed the seasonal differential geographically and temporally weighted regression (seasonal-difference GTWR, SD-GTWR) [35]. The latter authors applied the SD-GTWR model to data for hemorrhagic fever with renal syndromes from Hubei Province to show that the SD-GTWR model is superior to the ordinary GTWR model. The SD-GTWR model relied on the results of incremental spatial autocorrelation when balancing the roles of space and time. Data from the Zhejiang coast (China) from 2012 through 2016 was employed by Du et al. [36] to propose a geographical and periodic time-weighted regression model (GcTWR) that unifies spatial distance and temporal distance. The results confirmed that the seasonal effects on coastal areas are related to an interannual effect.

Although the GWR model performs well in the spatial interpolation, precipitation not only has continuity in space, but also has strong continuity in time. However, few kinds of research have been carried out on the spatial–temporal interpolation of precipitation. Fortunately, the application of the GTWR model in different fields has gradually become mature, and it is possible to introduce it into precipitation interpolation. Therefore, to better understand the temporal and spatial heterogeneity of precipitation, this paper interpolates the spatial distribution from 2006 through 2014 of the monthly and annual rainfall in Hubei Province, based on the GTWR model. Then, this work adjusts the spatial and temporal weight, according to the temporal characteristics of precipitation. The Gaussian kernel model is selected as the spatial weight, and the exponential function model is chosen as the temporal weight. Meanwhile, this work introduces the kriging model to eliminate the influence of residual spatial correlation on model fitting, which can improve the interpolation accuracy.

2. Materials and Methods

2.1. Methodology

2.1.1. Geographically and Temporally Weighted Regression Model

The GTWR model is expressed as Equation (1) [19,37]:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{n} \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i, i = 1, 2, \ldots, n \tag{1}$$

where (u_i, v_i, t_i) denotes the spatial–temporal coordinates of the observed location i, which contains the location and temporal information; k denotes the index $k = 1, 2 \ldots n$. $\beta_k(u_i, v_i, t_i)$ represents a set of values for the number n of parameters at point i, and ε_i represents the random error of the predicted variable y_i. The estimates of $\beta_k(u_i, v_i, t_i)$ are given by Formula (2):

$$\widehat{\beta}_k(u_i, v_i, t_i) = \left(X^T W(u_i, v_i, t_i)X\right)^{-1} X^T W(u_i, v_i, t_i)y \tag{2}$$

The spatial–temporal weight matrix $W(u_i, v_i, t_i)$ is based on the definition of the spatial–temporal distance and its decay functions. Generally, the weight functions include the distance threshold method, distance inverse ratio method, Gaussian (Gauss) function method, and double square root (bi-square) kernel function method. The Gaussian and bi-square kernel function methods are commonly used in the GTWR model [38]. The Gaussian kernel function is given by Equation (3):

$$w_{ij} = \exp(-(d_{ij}^{ST}/h^{ST})^2) \tag{3}$$

where the element w_{ij} of the weighting matrix is determined by the spatial–temporal distance d_{ij}^{ST} and spatial–temporal bandwidth h^{ST}. The GTWR model sets the spatial–temporal distance d_{ij}^{ST} as a function of the temporal distance d_{ij}^T and the spatial distance, d_{ij}^S as expressed in Equation (4):

$$\left(d_{ij}^{ST}\right)^2 = \lambda\left(d_{ij}^S\right)^2 + \mu\left(d_{ij}^T\right)^2 = \lambda\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \mu\left(t_i - t_j\right)^2 \tag{4}$$

λ and μ denote the spatial distance factor and the temporal distance factor, respectively, which balance the effect of the space and time dimension on parameter estimation. The w_{ij} element of the spatial–temporal weight matrix is expressed by Equation (5):

$$
\begin{aligned}
w_{ij} &= \exp\left\{-\left(\frac{\lambda\left[\left(u_i-u_j\right)^2+\left(v_i-v_j\right)^2\right]+\mu\left(t_i-t_j\right)^2}{(h^{ST})^2}\right)\right\} \\
&= \exp\left\{-\left[\frac{\left(u_i-u_j\right)^2+\left(v_i-v_j\right)^2}{(h^S)^2}+\frac{\left(t_i-t_j\right)^2}{(h^T)^2}\right]\right\} \\
&= \exp\left\{-\left[\frac{\left(d_{ij}^S\right)^2}{(h^S)^2}+\frac{\left(d_{ij}^T\right)^2}{(h^T)^2}\right]\right\} \\
&= \exp\left\{-\left[\frac{\left(d_{ij}^S\right)^2}{(h^S)^2}\right]\right\} \times \exp\left\{-\left[\frac{\left(d_{ij}^T\right)^2}{(h^T)^2}\right]\right\} \\
&= w_{ij}^S \times w_{ij}^T
\end{aligned}
\tag{5}
$$

It is seen in Equation (5) that the spatial–temporal element or kernel function w_{ij} equals the spatial kernel function w_{ij}^S multiplied by the temporal kernel function w_{ij}^T. h^S and h^T are the spatial and temporal bandwidths, respectively. One can determine the weight of the observed variable at a given

location to the regressed variable at the same location at a specific time by the spatial–temporal kernel function. The spatial bandwidth b_S and temporal bandwidth b_T are decided by cross-validation (CV), and they are obtained by minimizing the expression on the right-hand side of Equation (6):

$$CV(b_S, b_T) = \frac{1}{n} \sum_{i=1}^{n} \left[y_i - \widehat{y}_{\neq i}(b_S, b_T) \right]^2 \tag{6}$$

when $CV(\lambda, \mu)$ is minimized, it yields the optimal spatial bandwidth and temporal bandwidth of the model. A method for determining spatial and temporal bandwidth in steps was proposed by Fotheringham et al. [21]. The principle of this method is that GTWR can be regarded as GWR for a period of time. In each time period of data, the spatial bandwidth is determined by the GWR model by minimizing $CV(b_S)$, and then determining the time bandwidth that minimizes CV $(b_{S1}, b_{S2}, \ldots, b_{Sn}, b_T)$.

One obtains the optimal spatial bandwidth and temporal bandwidth when the CV is minimized. This paper relies on a step-by-step approach [8] to calculate the bandwidth. This approach considers the data at a specific time or a period at first. The GTWR became similar to the GWR model in this manner. One minimize the spatial data $CV(b_S)$ of each period, and then obtain the appropriate time bandwidth by minimizing the CV $(b_{S1}, b_{S2}, \ldots, b_{Sn}, b_T)$.

2.1.2. Comparison Models with Different Spatial–Temporal Parameters

The spatial–temporal distance calculation method of the GTWR model is similar to the extension method, which adds the temporal distance as the third dimension to the distance calculation. Previous studies have generally neglected the shortcomings of the extension method, whereby the uncertainty of the units (say, m or km for spatial units) introduces uncertainty of the spatial–temporal interpolation results when calculating the spatial–temporal distance [13]. The calculated results differ substantially depending on the adopted units, such as the unit of spatial distance being in meters or kilometers, and the time unit being the year, month, day, minute, or second. This paper chooses km as the spatial distance unit, whereas the timescale may be annual, monthly, and daily, according to the experimental data scale.

The temporal weighting formulas may be inconsistent with the spatial weighting formulas whenever the spatial–temporal weight is decomposed into the product of the spatial weight and the temporal weight. This work proposes a time weight in the form of an exponential function, as shown in Formula (7), in agreement with the last line on the right-hand side of Equation (5):

$$w_{ij}^T = \exp(-d_{ij}^T / h^T) \tag{7}$$

Furthermore, the distribution of variables or objects in the time dimension is not entirely governed by the previously cited first law of geography. There are apparent cyclical changes in the four seasons of the year. The physical characteristics of the climate at a location differ in the winter compared with summer, but they have statistical similarities for the same season in different years. Concerning monthly precipitation it is known that there may be statistical similarities between a given month's precipitation and non-adjacent month precipitation in the same quarter. Thus, it is necessary to improve the calculation method of the spatial–temporal distance to capture such similarities. The temporal distribution of precipitation exhibits periodicity, therefore, this paper relies on the sinusoidal function to calculate the temporal distance. It selects the exponential function model as a temporal weight formula. The calculation formula of the periodic temporal distance is given by Equation (8), where T denotes the period of the function:

$$d_{ij}^T = \sin\left(\frac{(t_i - t_j)\pi}{T}\right)(t_i - t_j) \tag{8}$$

2.1.3. Geographically and Temporally Weighted Regression Kriging

It is a spatial interpolation method based on geostatistics of the kriging method which fully considers the characteristics of the spatial variability of the sample points. It has the advantages of strong applicability and high prediction accuracy, and is currently most widely used in the fields of meteorology, ecology, and soil. The kriging algorithm can achieve optimal linear unbiased values, improving the accuracy of estimation to a certain extent. Moreover, for regionalized variables, it reveals its spatial structure well.

The GTWR model embeds temporal information and geographic location into the model, making full use of the spatial and temporal characteristics of data, and has a useful application in regional regression analysis. Geographically and temporally weighted regression kriging (GTWRK) is a hybrid method based on the GTWR model. First, we use the GTWR method to establish the regression relationship between precipitation and auxiliary information. Second, we use the kriging method to interpolate the residuals ε of the GTWR model. Finally, we add the interpolation result of residuals and the GTWR regression estimation value to obtain the GTWRK estimation result. Therefore, the GTWRK method considers the relationship between precipitation and influencing factors and the spatial autocorrelation of precipitation. The GTWR model is given by Equation (9):

$$\hat{y}_{GTWRK}(u_i, v_i, t_i) = \hat{y}_{GTWR}(u_i, v_i, t_i) + \hat{\varepsilon}_{OK}(u_i, v_i, t_i) \tag{9}$$

where: $\hat{y}_{GTWRK}(u_i, v_i, t_i)$ $\hat{y}_{GTWRK}(u_i, v_i, t_i)$ is the estimated value of GTWRK; $\hat{y}_{GTWR}(u_i, v_i, t_i)$ denotes the estimated value of GTWR; and $\hat{\varepsilon}_{OK}(u_i, v_i, t_i)$ represents the residual interpolation result of GTWR regression obtained by ordinary kriging (OK) interpolation. The variogram must be selected when the kriging method is used to interpolate the residuals. This work chooses an exponential variogram for ordinary kriging interpolation based on exploratory analysis of the precipitation data.

The flow chart of the GTWRK model is shown in Figure 1.
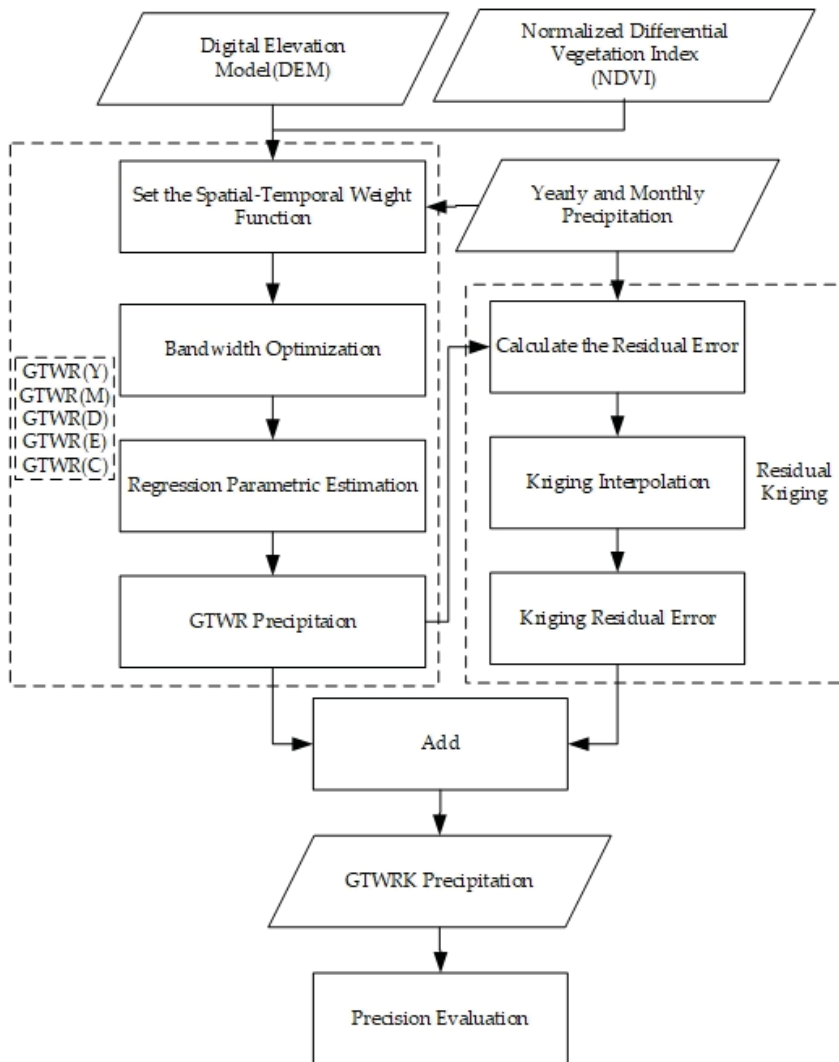
Figure 1. The flow chart of the geographically and temporally weighted regression kriging (GTWRK) model.

2.1.4. Precision Evaluation

Because of the large amount of data generated in the interpolation comparison, it would be burdensome to display all the results. The results corresponding to precipitation fitting in July 2008 and May 2013 are representative of the monthly scale data, and the results of precipitation fitting in 2010 and 2012 are representative of the annual scale data. The evaluation of the interpolation models relied on several performance indices, namely the mean absolute error (MAE), mean relative error (MRE), and the root mean square error (RMSE). The smaller the values of MAE, MRE, and RMSE, the better the interpolation effect.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i| \tag{10}$$

$$MRE = \frac{1}{n} \sum_{i=1}^{n} \frac{\left|Y_i - \hat{Y}_i\right|}{Y_i} \qquad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2} \qquad (12)$$

in which n denotes the sample size, Y_i represents the ith a sample value, $\hat{Y}_i$ denots the sample estimates.

2.2. Study Area and Data

Hubei Province is located between northern latitudes 29°05′ and 33°20′ and eastern latitudes 108°21′ and 116°07′ within China. The province covers an area of 185,900 square kilometers, and includes a variety of topographical regions, including mountains, plains, and transitional topographic zones. The topography of Hubei Province exhibits the highest elevation of 3090 m in its western region, and lowest towards its eastern area, with a lowest elevation of −142 m [39]. Hubei Province lies within a north–south transitional climatic zone, belonging to the northeastern Asian monsoonal region, except for the higher elevations of the western mountainous areas. Most of the province features a subtropical monsoonal, humid climate. The warm temperature coincides with the rainy period, providing abundant rainfall to support agroforest production [39]. The spatial distribution of meteorological sites and the digital elevation model (DEM) are shown in Figure 2.
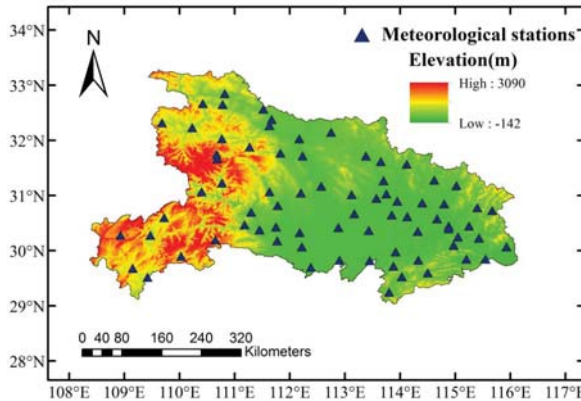
Figure 2. The spatial distribution of meteorological sites and digital elevation model (DEM) data representation.

The Hubei Meteorological Bureau provided annual and monthly precipitation data from 75 meteorological stations from 2006 through 2014 in Hubei Province. Figure 3 displays the average monthly precipitation distribution for the period of record.
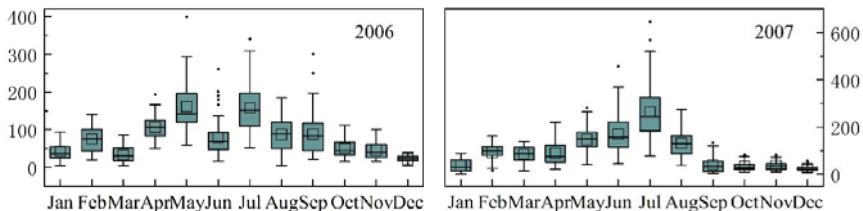
Figure 3. *Cont.*
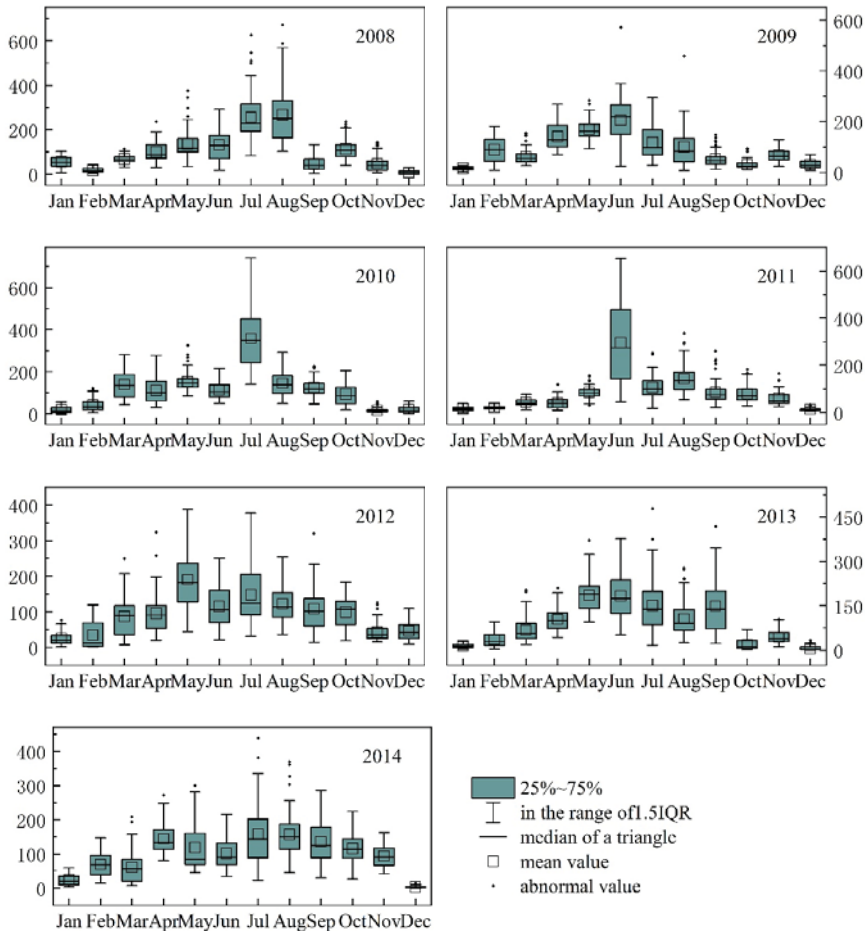
Figure 3. The monthly precipitation distribution from 2006 to 2014.

Many factors influence the spatial distribution of precipitation. Previous studies have demonstrated that topographic characteristics are the main controlling factor [40,41]. The Normalized Differential Vegetation Index (NDVI) is influenced by precipitation, and vice versa [42,43]. This work employs the Digital Elevation Model (DEM) Shuttle Radar Topography Mission (SRTM) data with 90 m resolution. The NDVI data are the MODND1M for China with 500 m resolution NDVI monthly synthesis products. All of the above data are available from the Geospatial Data Cloud website (http://www.gscloud.cn).

2.3. Data Preprocessing

All data were consolidated under the WGS-84 geographic coordinate system. The grid resolution employed was $0.1° \times 0.1°$. Precipitation constitutes the dependent variable; longitude, latitude, DEM, and the NDVI are the independent variables. The timescale of precipitation data is divided into monthly and annual categories. The monthly data represent the time series data from January 2006 through December 2014. The yearly data denote the time series data from 2006 through 2014.

The interpolation accuracy of the GTWR models was evaluated by cross-verification (CV). The 75 meteorological sites were randomly divided into 60 sites (80% of the total number of sites) as the modeling set for interpolation; the remaining 15 sites (20% of the total number of sites) were chosen as the validation set, which was used to evaluate the models' accuracies. This paper employs the GTWR models corresponding to several timescales listed in Table 1.

Table 1. Listing of the geographically and temporally weighted regression (GTWR) models.

Model	Spatial Weight	Time Weight	Timescale	Calculation Method	Residual Processing
GTWR(Y)	Gaussian	Gaussian	Year	Subtraction	None
GTWR(M)	Gaussian	Gaussian	Month	Subtraction	None
GTWR(D)	Gaussian	Gaussian	Day	Subtraction	None
GTWR(E)	Gaussian	Exponential	Day	Subtraction	None
GTWR(C)	Gaussian	Exponential	Day	Sinusoidal	None
GTWRK	Gaussian	Exponential	Day	Subtraction	Kriging interpolation

3. Results and Analysis

3.1. Monthly Data

The fitting interpolation distribution maps for July 2008 and May 2013 are shown in Figures 4 and 5, respectively.
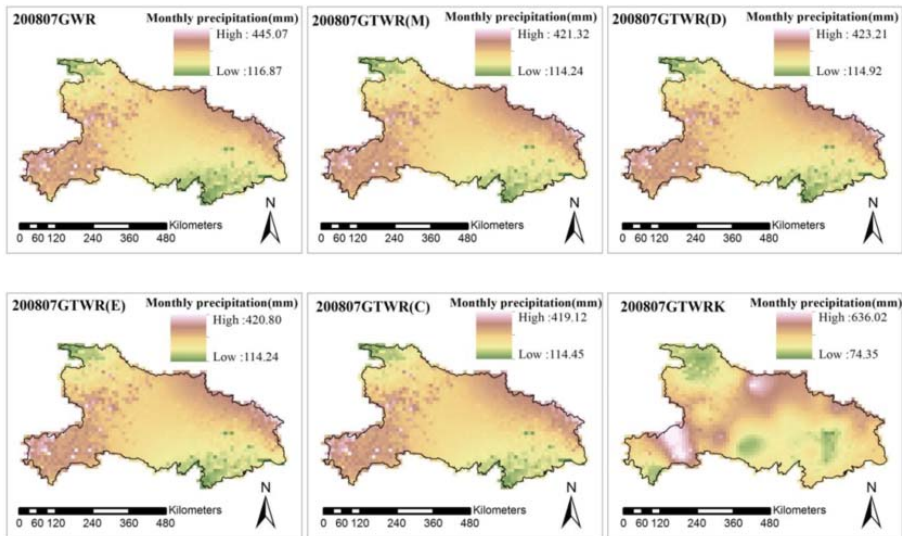
Figure 4. The fitting interpolation distribution map for geographically weighted regression (GWR), GTWR(M), GTWR(D), GTWR(E), and GTWR(C) corresponding to July 2008.
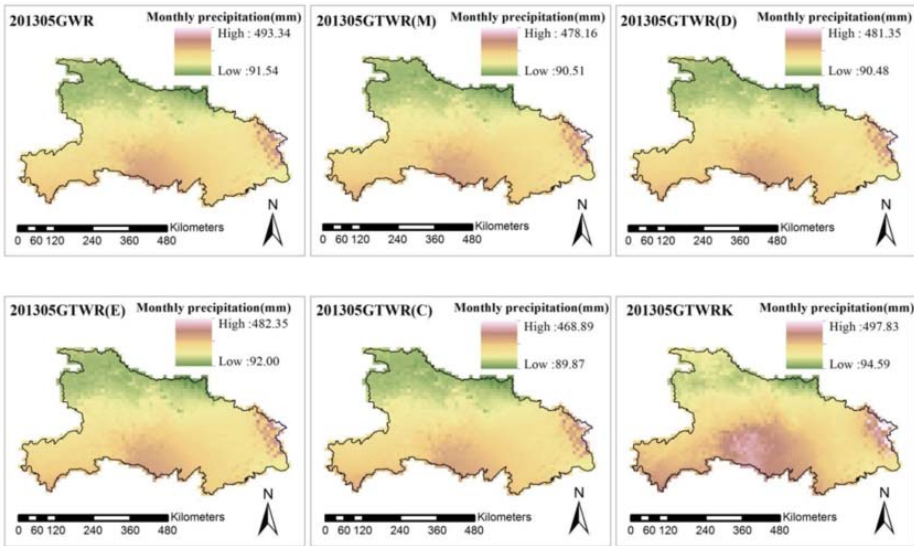
Figure 5. The interpolation distribution maps for GWR, GTWR(M), GTWR(D), GTWR(E), and GTWR(C) corresponding to May 2013.

The average error maps of MAE, MRE, and RMSE for monthly scale data are shown in Figure 6.
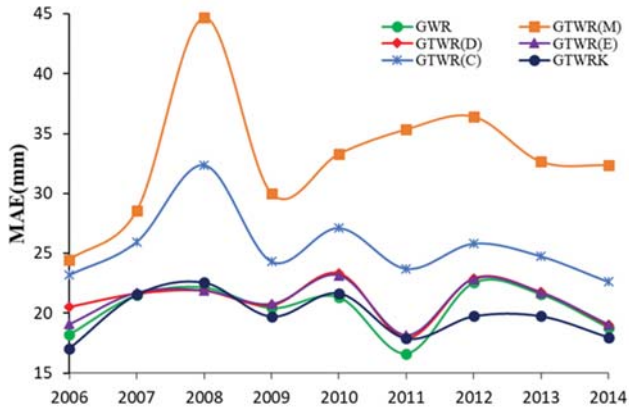
(**a**) MAE distribution chart of monthly scale data

Figure 6. *Cont.*
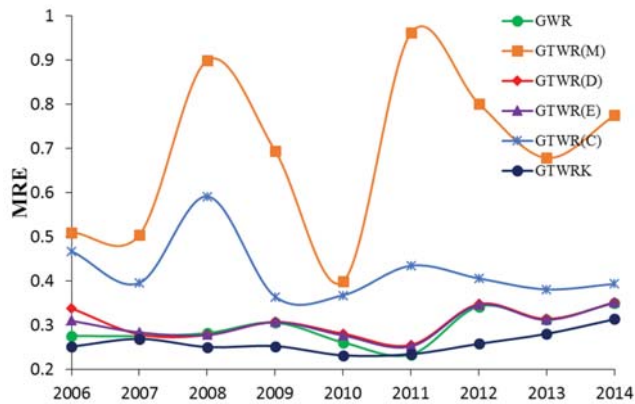
(**b**) MRE distribution chart of monthly scale data
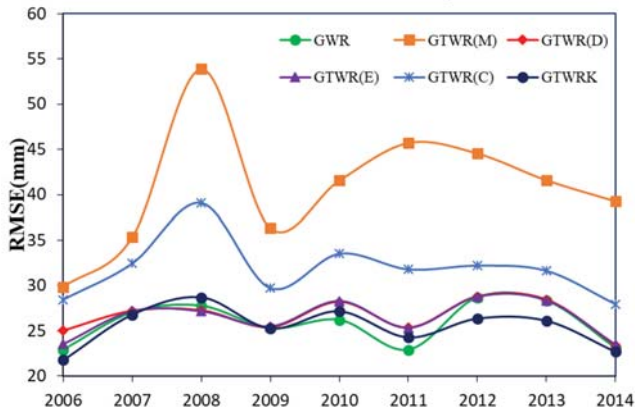
(**c**) RMSE distribution chart of monthly scale data

Figure 6. The average error maps for monthly scale data from 2006 through 2014. (**a**–**c**) represent the average distribution chart of mean average error (MAE), mean root error (MRE), and root mean square error (RMSE), respectively.

It can be seen in Figure 6 and Table 2 that the optimal timescale of the GTWR model is daily for the monthly scale data. The MAE, MRE, and RMSE decrease by 36%, 56%, and 35%, respectively, when choosing GTWR(D) instead of GTWR(M). The GTWR(E) improves the accuracy of the results compared to GTWR(D), by reducing the MAE, MRE, and RMSE by 0.7%, 1.1%, and 0.6%, respectively. The fitting accuracy of GTWR(E) and GWR are similar, with a difference of about 3% for the monthly scale data results shown in Table 2. The GTWR(C) has a lower accuracy compared with GWR, and increased MAE, MRE, and RMSE by 25%, 45%, and 24%. The GTWRK has the highest accuracy compared with GWR, and decreased MAE, MRE, and RMSE by 3%, 10%, and 1%, respectively.

Figure 7 shows the MAE, MRE, and RMSE of monthly GWR and GTWRK models in different seasons. According to the character, the fitting accuracy of the GTWRK model is higher than that of the GWR model as a whole, especially in spring, autumn, and winter.

Table 2. The average MAE, MRE, and RMSE for the monthly scale data.

Model	MAE (mm)	MRE	RMSE (mm)
GWR	20.36	0.29	25.80
GTWR(M)	33.08	0.69	40.90
GTWR(D)	21.09	0.31	26.57
GTWR(E)	20.92	0.30	26.40
GTWR(C)	25.54	0.42	31.88
GTWRK	19.77	0.26	25.47

(**a**) Comparison of evaluation indicators of GWR and GTWRK models in spring

(**b**) Comparison of evaluation indicators of GWR and GTWRK models in summer

(**c**) Comparison of evaluation indicators of GWR and GTWRK models in autumn

(**d**) Comparison of evaluation indicators of GWR and GTWRK models in winter
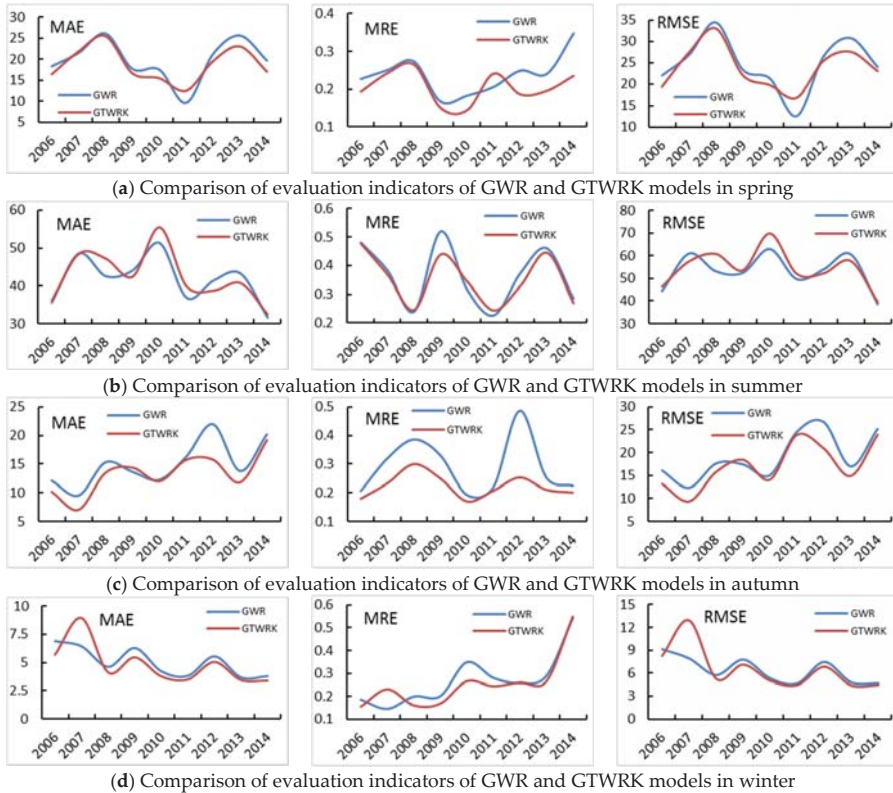
Figure 7. Applicability of GWR and GTWRK models in different seasons. (**a–d**) represent the distribution chart of MAE (Units: mm), MRE, and RMSE (Units: mm) in the four seasons, respectively.

3.2. Annual Data

The interpolation distribution maps corresponding to 2010 and 2012 are shown in Figures 8 and 9, respectively.
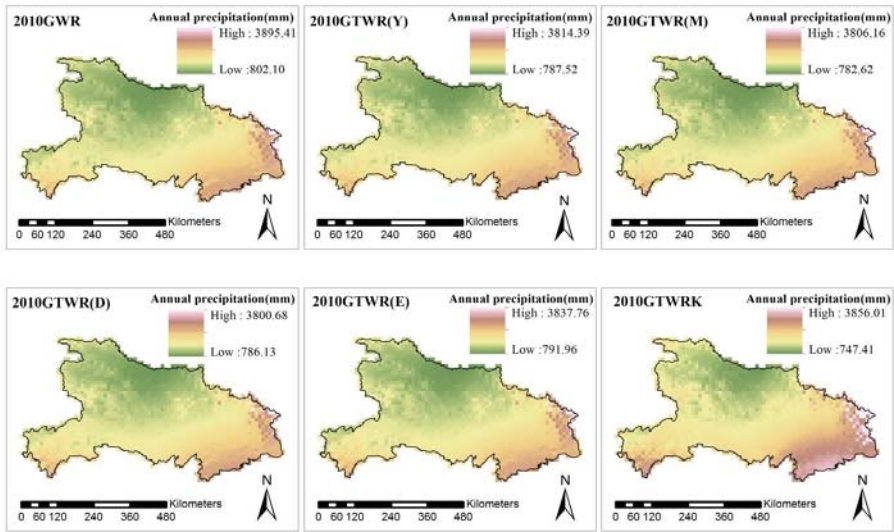
Figure 8. The interpolation distribution maps for GWR, GTWR(Y), GTWR(M), GTWR(D), and GTWR(E) corresponding to 2010.
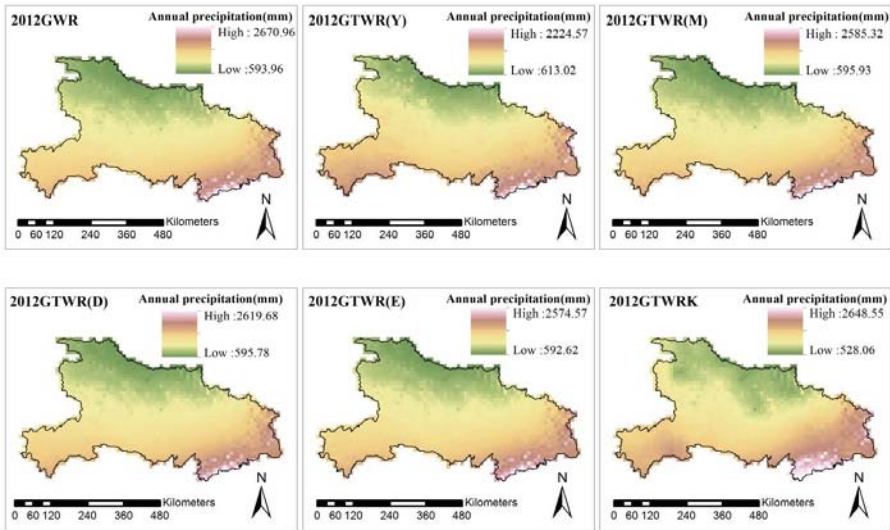
Figure 9. The interpolation distribution maps for GWR, GTWR(Y), GTWR(M), GTWR(D), and GTWR(E) corresponding to 2012.

It is seen in Figure 10 and Table 3 that the optimal timescale of the GTWR model is daily for the annual scale data. The MAE, MRE, RMSE decrease by 13%, 15%, and 14%, respectively, when choosing the GTWR(Y) over the GTWR(D). In the results, the accuracy error of GTWR(E) can be reduced by about 0.2%, compared to GTWR(D). When the GTWRK is compared to the GWR model, MAE, MRE, and RMSE decrease by 3%, 10%, and 5%, respectively.
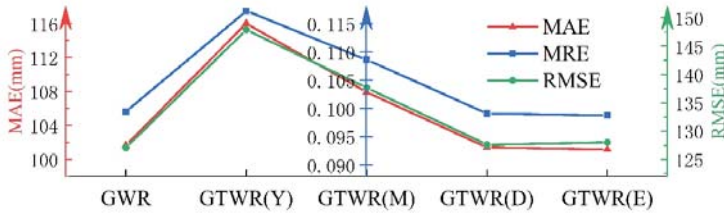
Figure 10. The average error for annual data.

Table 3. The average MAE, MRE, and RMSE for the annual data.

Model	MAE (mm)	MRE	RMSE (mm)
GWR	101.64	0.10	127.08
GTWR(Y)	116.04	0.12	147.94
GTWR(M)	107.93	0.11	137.69
GTWR(D)	101.41	0.10	127.57
GTWR(E)	101.17	0.10	128.01
GTWRK	98.27	0.09	120.60

4. Discussion

The performances of GWR and GTWR models are relatively similar at the annual and monthly scales according to the reductions in average MAE, MRE, and RMSE. However the application of the GTWR model is not effective with respect to all the indicators, which reflects the extreme uncertainties of rainfall in time.

It is preferable to use a daily time scale to calculate spatial weights. Precipitation is formed by precipitation processes, and the duration of each process is usually several days. Therefore, as the timescale goes from yearly to daily the fine temporal details of precipitation gradually become prominent. Figure 11 shows the precipitation process in August 2008. The average monthly precipitation stands for the rainfall of the whole month. Consequently, the performance of the GTWR(D) model with a daily time scale is much better than the GTWR(M) and GTWR(Y) models' at the monthly (Table 2) and annual (Table 3) scales.

The periodicity of precipitation has some impact on improving the accuracy of interpolation. As shown in Figure 6 shows the GTWR(C) model performs better than GTWR(M). However, the frequency and amplitude of each precipitation cycle are very different, as shown in Figures 3 and 11. The calculation of the periodic function needs further study.

The introduction of kriging is reasonable in the improvement of interpolation accuracy. Table 4 shows the normality test of residuals between the results of the GTWR model and the actual precipitation. The residuals of most months fitted normal distribution (significance > 0.05), while other months, such as May, June, September, and December, fitted approximate normal distribution (kurtosis < 10 and skewness < 3). Compared with the GTWR model, the GTWRK model performed well in terms of average MAE, MRE, and RMSE (Tables 2 and 3). The GTWRK model produces a more accurate spatio-temporal precipitation. The GTWRK model's performance varies performs through the seasons. This work represent summer as June, July, and August. It seen in Table 4, the residuals in these months all fitted the approximately normal distribution, which affects the accuracy of the GTWRK model to some extent.

(**a**) Average monthly precipitation from 2006 to 2014   (**b**) Monthly precipitation in 2008

(**c**) the daily precipitation in August 2008
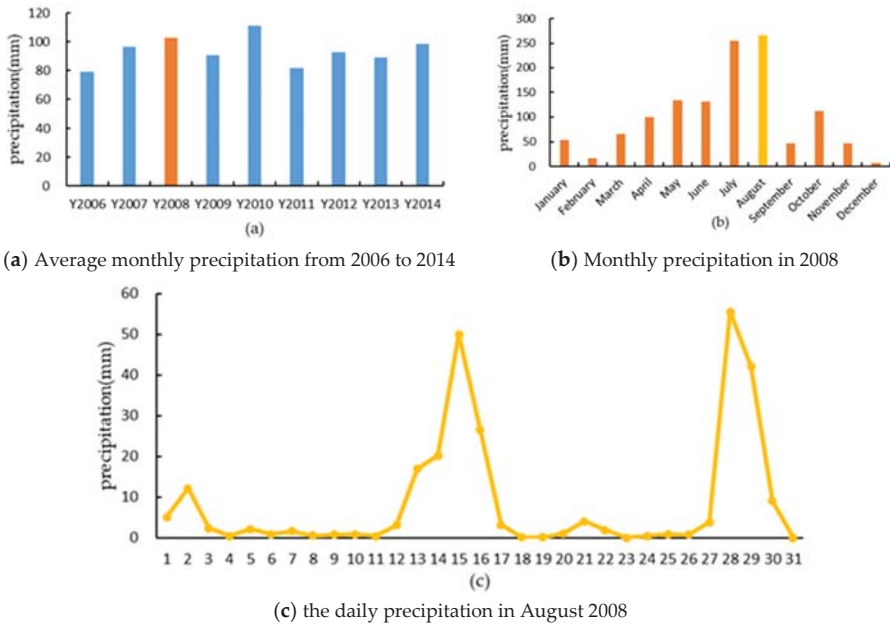
Figure 11. The average precipitation of Hubei.

Table 4. The normality test of residuals between GTWR and the actual precipitation.

Month	Kolmogorov-Smirnova		Kurtosis	Skewness	Average Precipitation (mm)
	df	Sig.			
January	60	0.200 *	0.674	0.553	53.99
February	60	0.200 *	−0.424	−0.01	16.29
March	60	0.200 *	1.228	0.448	65.99
April	60	0.200 *	5.968	1.661	99.41
May	60	0.005	3.979	1.802	134.22
June	60	0.006	−0.139	0.844	130.83
July	60	0.059	0.95	0.939	255.36
August	60	0.054	0.922	0.886	266.10
September	60	0.039	0.342	0.626	46.81
October	60	0.200 *	0.965	0.273	111.70
November	60	0.200 *	0.471	−0.08	46.72
December	60	0	0.709	0.386	7.01

* This is a lower bound of the true significance; a. Lilliefors significance correction.

5. Conclusions

In this work, a GTWRK model combined with the GTWR and kriging model was introduced to interpolate the spatial distribution of monthly and annual precipitation from 2006 through 2014 in Hubei Province. The main conclusions are as follows:

(1) GTWRK obtains a better average interpolation accuracy, compared to the GWR model. In the comparison between the GTWRK and GWR, the MAE decreased from 101.64 to 98.27. Consequently, we conclude that it is an improvement to extend GTWR with kriging.

(2) The optimal timescale for interpolating precipitation data with the GTWR model is daily. The fitting accuracy is improved when the timescale is converted from yearly to daily. Compared with the GTWR(M) model, the average MAE, MRE, and RMSE of the monthly scale data decreased

by 36%, 56%, and 35%, respectively, when using daily data. The same indices for the annual data reduced by 13%, 15%, and 14% when using daily data, respectively.

(3)　The temporal weight based on an exponential function improved the GTWR model at the monthly and annual data. It reduced the accuracy difference of the monthly scale between GTWR and GWR by about 3%. For the yearly scale data, the years with improved accuracy account for about 55%. Especially in 2008, 2009, 2010, 2011, and 2013, the accuracy was improved significantly. Meanwhile, the GTWRK improves the accuracy as measured by the MAE, MRE, and RMSE by 3%, 10%, and 1%, respectively, of monthly precipitation prediction, and by 3%, 10%, and 5%, respectively, of annual precipitation predictions.

(4)　The proposed model could be applied to manage similar phenomena with a large historical dataset. Meanwhile, the GTWR model takes into account the spatial and temporal heterogeneity of precipitation and produces better estimates of the residuals.

(5)　This work explored the annual, monthly, and daily scales to adjust the optimal time scale, while other time scales should be explored in future work. Additionally, the influence of the periodic characteristics of precipitation on the GTWR model needs further study.

Author Contributions: W.Z. proposed the theme and method of the article, conducted the initial analysis, provided the financial support for the project leading to this publication and provided the creation of the revision and published work; D.L. conducted the research and investigation process, performed the experiment of the thesis, and write the original draft; S.Z. provided and processed the source of the data, and participated the investigation process; S.L. provided the presentation of the revision and published work; H.A.L. provided critical review, data interpretation and presentation of results, and English editing in the revision and publication stages; W.L. took part in the presentation of the published work on revision and publication stages. All authors have read and agreed to the published version of the manuscript.

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

References

1.　Bhattacharya, A.; Adhikari, A.; Maitra, A. Multi-technique observations on precipitation and other related phenomena during cyclone Aila at a tropical location. *Int. J. Remote Sens.* **2013**, *34*, 1965–1980. [CrossRef]
2.　Goovaerts, P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **2000**, *228*, 113–129. [CrossRef]
3.　Vicente-Serrano, S.M.; Beguería, S.; López-Moreno, J.I. A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Clim.* **2010**, *23*, 1696–1718. [CrossRef]
4.　Plouffe, C.C.F.; Robertson, C.; Chandrapala, L. Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: A case study of Sri Lanka. *Environ. Model. Softw.* **2015**, *67*, 57–71. [CrossRef]
5.　Collischonn, B.; Collischonn, W.; Tucci, C.E.M. Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *J. Hydrol.* **2008**, *360*, 207–216. [CrossRef]
6.　Lin, Z.; Mo, X.; Li, H.; Li, H. Comparison of Three Spatial Interpolation Methods for Climate Variables in China. *Acta Geogr. Sin.* **2002**, *57*, 47–56.
7.　Feng, Z.; Yang, Y.; Ding, X.; Lin, Z. Optimization of the spatial interpolation methods for climate resources. *Geogr. Res.* **2004**, *23*, 357–364.
8.　Liu, J.-F.; Chen, R.-S.; Han, C.-T.; Tan, C.-P. Evaluating TRMM multi-satellite precipitation analysis using gauge precipitation and MODIS snow-cover products. *Adv. Water Sci.* **2010**, *21*, 343–348.
9.　Zhu, L.; Huang, J. Comparis on of spatial interpolation m ethod for precipitation of mountain areas in county scale. *Trans. CSAE* **2007**, *23*, 80–85.
10.　Li, C.; Chen, L.; Wang, Y. Research on spatial interpolation of rainfall distribution-A case study of Idaho State in the USA. *Miner. Resour. Geol.* **2007**, *21*, 684–687.
11.　Lam, N.S.-N. Spatial Interpolation Methods: A Review. *Am. Cartogr.* **1983**, *10*, 129–150. [CrossRef]
12.　Li, L.; Revesz, P. Interpolation methods for spatio-temporal geographic data. *Comput. Environ. Urban. Syst.* **2004**, *28*, 201–227. [CrossRef]

13. Peng, S. Developments of Spatio-temporal Interpolation Methods for Meteorological Elements. Master's Thesis, Central South University, Changsha, China, 2010.

14. Li, S.; Shu, H.; Xu, Z. Interpolation of temperature based on spatial-temporal Kriging. *Geomatics Inf. Sci. Wuhan Univ.* **2012**, *37*, 237–241.

15. Lu, Y. Spatio-Temporal Cokriging Interpolation for Air Pollution Index Analysis. Master's Thesis, Chinese Academy of Surveying & Mapping, Beijing, China, 2018.

16. Fotheringham, A.S.; Brunsdon, C.; Charlton, M.E. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298.

17. Hu, Q.; Li, Z.; Wang, L.; Huang, Y.; Wang, Y.; Li, L. Rainfall Spatial Estimations: A Review from Spatial Interpolation to Multi-Source Data Merging. *Water* **2019**, *11*, 579. [CrossRef]

18. Wang, M.; He, G.; Zhang, Z.; Wang, G.; Zhang, Z.; Cao, X.; Wu, Z.; Liu, X. Comparison of Spatial Interpolation and Regression Analysis Models for an Estimation of Monthly Near Surface Air Temperature in China. *Remote. Sens.* **2017**, *9*, 1278. [CrossRef]

19. Huang, B.; Wu, B.; Barry, M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 383–401. [CrossRef]

20. Wu, B.; Li, R.; Huang, B. A geographically and temporally weighted autoregressive model with application to housing prices. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1186–1204. [CrossRef]

21. Fotheringham, A.S.; Crespo, R.; Yao, J. Geographical and Temporal Weighted Regression (GTWR). *Geogr. Anal.* **2015**, *47*, 431–452. [CrossRef]

22. Xiao, H.; Yi, D. Empirical study of carbon emission drivers based on Geographically time weighted regression model. *Stat. Inf. Forum* **2014**, *29*, 83–89.

23. Liu, J.; Yang, Y.; Xu, S.; Zhao, Y.; Wang, Y.; Zhang, F. A geographically temporal weighted regression approach with travel distance for house price estimation. *Entropy* **2016**, *18*, 303. [CrossRef]

24. Song, C.; Kwan, M.P.; Zhu, J. Modeling fire occurrence at the city scale: A comparison between geographically weighted regression and global linear regression. *Int. J. Environ. Res. Public Health* **2017**, *14*, 396. [CrossRef] [PubMed]

25. Zhou, Y.; Wu, L.; Zhang, Y.; Shen, Y.; Qin, K.; Bai, Y. A Geographically and Temporally Weighted Regression Model for Ground-Level PM2.5 Estimation from Satellite-Derived 500 m Resolution AOD. *Remote Sens.* **2016**, *8*, 262.

26. Qin, K.; Rao, L.; Xu, J.; Bai, Y.; Zou, J.; Hao, N.; Li, S.; Yu, C. Estimating ground level NO2 concentrations over central-eastern China using a satellite-based geographically and temporally weighted regression model. *Remote Sens.* **2017**, *9*, 950. [CrossRef]

27. Wei, Q.; Zhang, L.; Duan, W.; Zhen, Z. Global and Geographically and Temporally Weighted Regression Models for Modeling PM$_{2.5}$ in Heilongjiang, China from 2015 to 2018. *Int. J. Environ. Res. Public Health* **2019**, *16*, 5107. [CrossRef] [PubMed]

28. Brunsdon, C.; McClatchey, J.; Unwin, D.J. Spatial variations in the average rainfall-altitude relationship in Great Britain: An approach using geographically weighted regression. *Int. J. Climatol.* **2001**, *21*, 455–466. [CrossRef]

29. Chen, C.; Zhao, S.; Duan, Z.; Qin, Z. An Improved Spatial Downscaling Procedure for TRMM 3B43 Precipitation Product Using Geographically Weighted Regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4592–4604. [CrossRef]

30. Lv, A.; Zhou, L. A rainfall model based on a Geographically Weighted Regression algorithm for rainfall estimations over the arid Qaidam Basin in China. *Remote Sens.* **2016**, *8*, 311. [CrossRef]

31. Georganos, S.; Abdi, A.M.; Tenenbaum, D.E.; Kalogirou, S. Examining the NDVI-rainfall relationship in the semi-arid Sahel using geographically weighted regression. *J. Arid Environ.* **2017**, *146*, 64–74. [CrossRef]

32. Li, Y.; Xiong, L.; Yan, L.A. Geographically Weighted Regression Krigin Approach for TRMM-Rain Gauge Data Merging and its Application in Hydrological Forecasting. *Resour. Environ. Yangtze Basin* **2017**, *26*, 1359–1368.

33. Liu, J.; Zhao, Y.; Yang, Y.; Xu, S.; Zhang, F.; Zhang, X.; Shi, L.; Qiu, A. A mixed geographically and temporallyweighted regression: Exploring spatial-temporal variations from global and local perspectives. *Entropy* **2017**, *19*, 53. [CrossRef]

34. Tobler, W. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46* (Suppl. 46), 234–240. [CrossRef]

35. Ge, L.; Zhao, Y.; Sheng, Z.; Wang, N.; Zhou, K.; Mu, X.; Guo, L.; Wang, T.; Yang, Z.; Huo, X. Construction of a seasonal difference-geographically and temporally weighted regression (SD-GTWR) model and comparative analysis with GWR-based models for hemorrhagic fever with renal syndrome (HFRS) in Hubei Province (China). *Int. J. Environ. Res. Public Health* **2016**, *13*, 1062. [CrossRef] [PubMed]

36. Du, Z.; Wu, S.; Zhang, F.; Liu, R.; Zhou, Y. Extending geographically and temporally weighted regression to account for both spatiotemporal heterogeneity and seasonal variations in coastal seas. *Ecol. Inform.* **2018**, *43*, 185–199. [CrossRef]

37. Dong, F.; Wang, Y.; Zhang, X. Can Environmental Quality Improvement and Emission Reduction Targets Be Realized Simultaneously? Evidence from China and A Geographically and Temporally Weighted Regression Model. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2343. [CrossRef] [PubMed]

38. Qin, W. The Basic Theoretics and Application Research on Geographically Weighted Regression. Ph.D. Thesis, Tongji University, Shanghai, China, 2007.

39. Yuan, M. Dynamic Change of Vegetation and Phenology Response to Climate Change in Hubei Province. Master's Thesis, Wuhan University, Wuhan, China, 2017.

40. Spreen, W.C. A determination of the effect of topography upon precipitation. *Trans. Am. Geophys. Union* **1947**, *28*, 285–290. [CrossRef]

41. Smith, R.B. The influence of mountains on the atmosphere. *Adv. Geophys.* **1979**, *21*, 87–230.

42. Kang, L.; Di, L.; Shao, Y.; Yu, E.; Zhang, B.; Shrestha, R. Study of the NDVI-precipitation correlation stratified by crop type and soil permeability. In Proceedings of the 2013 2nd International Conference on Agro-Geoinformatics: Information for Sustainable Agriculture, Agro-Geoinformatics, Fairfax, VA, USA, 12–16 August 2013; pp. 194–199.

43. Feng, J.; Zhang, H.; Hu, X.; Shi, Q.; Zubaidai, M. Spatial non-stationarity characteristics of the impacts of precipitation and temperature on vegetation coverage index: A case study in Yili River Valley, Xinjiang. *Acta Ecol. Sin.* **2016**, *36*, 4626–4634.

RAINBOW: An Operational Oriented Combined IR-Algorithm

Leo Pio D'Adderio [1,*], Silvia Puca [2], Gianfranco Vulpiani [2], Marco Petracca [2], Paolo Sanò [1] and Stefano Dietrich [1]

[1] CNR-ISAC, Consiglio Nazionale delle Ricerche, Roma, Via del Fosso del Cavaliere 100, 00133 Roma, Italy; paolo.sano@artov.isac.cnr.it (P.S.); s.dietrich@isac.cnr.it (S.D.)
[2] Department of Civil Protection, Presidency of the Council of Ministers, Via Vitorchiano 2, 00189 Rome, Italy; Silvia.Puca@protezionecivile.it (S.P.); gianfranco.vulpiani@protezionecivile.it (G.V.); Marco.Petracca@protezionecivile.it (M.P.)
*   Correspondence: leopio.dadderio@artov.isac.cnr.it

Abstract: In this paper, precipitation estimates derived from the Italian ground radar network (IT GR) are used in conjunction with Spinning Enhanced Visible and InfraRed Imager (SEVIRI) measurements to develop an operational oriented algorithm (RAdar INfrared Blending algorithm for Operational Weather monitoring (RAINBOW)) able to provide precipitation pattern and intensity. The algorithm evaluates surface precipitation over five geographical boxes (in which the study area is divided). It is composed of two main modules that exploit a second-degree polynomial relationship between the SEVIRI brightness temperature at 10.8 μm $TB_{10.8}$ and the precipitation rate estimates from IT GR. These relationships are applied to each acquisition of SEVIRI in order to provide a surface precipitation map. The results, based on a number of case studies, show good performance of RAINBOW when it is compared with ground reference (precipitation rate map from interpolated rain gauge measurements), with high Probability of Detection (POD) and low False Alarm Ratio (FAR) values, especially for light to moderate precipitation range. At the same time, the mean error (ME) values are about 0 mmh^{-1}, while root mean square error (RMSE) is about 2 mmh^{-1}, highlighting a limited variability of the RAINBOW estimations. The precipitation retrievals from RAINBOW have been also compared with the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF) official microwave (MW)/infrared (IR) combined product (P-IN-SEVIRI). RAINBOW shows better performances than P-IN-SEVIRI, in terms of both detection and estimates of precipitation fields when they are compared to the ground reference. RAINBOW has been designed as an operational product, to provide complementary information to that of the national radar network where the IT GR coverage is absent, or the quality (expressed in terms of Quality Index (QI)) of the RAINBOW estimates is low. The aim of RAINBOW is to complement the radar and rain gauge network supporting the operational precipitation monitoring.

Keywords: remote sensing; precipitation; SEVIRI; ground radar

1. Introduction

Accurate precipitation measurements are essential for the validation of global climate models and for understanding the natural variability of the earth's climate. Moreover, rainfall monitoring can serve as an important element for risk management of severe precipitation events.

Although the importance of quantitative determination of rainfall is well recognized, reliable retrieval of precipitation is often difficult. First, precipitation represents one of the most difficult atmospheric variables to be accurately measured due to its high temporal and spatial variability.

Furthermore, the only instruments that guarantee direct measurements of precipitation are rain gauges and disdrometers. Both types of instruments, although, have a quite high temporal resolution, and provide point-like measurements, ensuring a low spatial resolution. On the other hand, ground-based radars provide measurements of rainfall with a relatively high spatial and temporal resolution. Although they represent a valuable source of information, they provide an indirect measurement of precipitation. In addition, radar observations are affected by several uncertainty sources, including miscalibration, ground clutter, beam blocking, attenuation, Wireless Local Area Network (W-LAN) interferences [1–4].

Space-borne monitoring of clouds and precipitation all around the globe has been gaining growing interest from the international scientific community as a primary contribution to the improvement of global precipitation measurement and to the determination and detection of the global climatic changes. Most of the space-borne monitoring systems take advantage of passive instrumentation, (e.g. radiometers), using both infrared (IR) and microwave (MW) emissions to retrieve cloud properties and precipitation estimation. However, it is difficult to establish an exact quantitative relationship between surface rain rate and the cloud physical quantities (e.g., brightness temperatures) measured by the various sensors [5–8]).

IR-based estimates of rainfall exploit the sensitivity of the IR measurements to the uppermost layers of clouds, but the measured cloud-top brightness temperatures do not provide sufficient information to retrieve the actual intensity of surface rainfall with high reliability. However, the relevance of IR estimates lie in the wide coverage of the earth at relatively high spatial and temporal resolution provided by geosynchronous satellites [9–13]), being IR sensors, mainly mounted on geostationary (GEO) satellites (e.g., the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard of Meteosat Second Generation (MSG) and the Geostationary Operational Environmental Satellite (GOES) Imagers).

However, rainfall estimates based on IR and VIS measurements are constantly evolving thanks also to the improved performance of the sensors. In this regard, it should be noted that the IR and VIS based rainfall retrievals have obtained an important improvement by the exploitation of optical and microphysical clouds parameters (e.g., optical thickness, particle radius), thanks to the higher enhanced spectral resolution of the new generation of geostationary sensors (e.g., MSG SEVIRI and GOES Imagers) [14–18]. In addition, the use of optical and microphysical cloud parameters, the use of classification schemes of convective and stratiform precipitation areas has also contributed to improving the accuracy of rainfall estimates [18,19]. Therefore, while the cloud-top temperature is a primary reference to detect deep convection and precipitation, the use of microphysics parameters and of the cloud classification schemes helps to solve the ambiguities in the retrieval and to identify more accurately the rainy area at the ground [20]. It is also worth mentioning that the combined use of both IR and VIS radiation to provide meteorological products supporting nowcasting activities has been widely studied in the EUMETSAT program—Satellite Application Facilities on Support to Nowcasting and Very Short Range Forecasting (NWC SAF) [20–22]. Furthermore, significant progresses are being made in the field of hyperspectral IR detection and substantial impacts are expected on the Numerical Weather Prediction (NWP) [23–25].

On the other hand, MW-based observations have the great advantage of providing a more direct measurement of the precipitation due to the ability of MW radiation to penetrate precipitating clouds and interact with its liquid and ice hydrometeors [26–30]). At the same time, they suffer of the insufficient temporal frequency of Low Earth Orbit (LEO) satellite overpasses (which carry MW instruments), with respect to the high variability of the precipitation in time and space.

To reduce the evidenced limitations and obtain satisfactory precipitation measurements in terms of accuracy, spatial, and temporal resolution, researchers have increasingly moved to using combinations of sensors. The joint use of MW and IR measurements has long been recognized as very effective as it combines the accuracy of the instantaneous MW data and the repetition and coverage characteristics of the IR geostationary measurements [12,31–34]).

The higher number of LEO-GEO satellites orbiting around the globe has made available a significant amount of precipitation estimates. The availability of these estimates are useful to build accurate and reliable multi-satellite datasets. The goal is to provide products with the best short-range estimates, called High Resolution Precipitation Products (HRPP). The Tropical Rainfall Measuring Mission's (TRMM) Multisatellite Precipitation Analysis (TMPA) was produced according to this line, since it combines precipitation estimates from multiple satellites, as well as from rain gauges, where feasible, to generate rainfall data [35,36].

The Climate Prediction Center morphing method (CMORPH) uses motion vectors from dynamic GEO-IR images to fill the temporal gaps between two available Passive Microwave (PMW) rainfall estimates [37]. The Japanese Global Precipitation Measurement (GPM) standard product Global Satellite Mapping of Precipitation (GSMaP) is a PMW–IR precipitation product. The algorithm integrates PMW data with infrared radiometer data to achieve high temporal and spatial resolution global precipitation estimates [38]. The National Oceanic and Atmospheric Administration (NOAA) Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR) algorithm estimates rainfall at a fine temporal resolution using PMW (SSM/I—-Special Sensor Microwave/Imager) and GEO (GOES) satellites. It uses SSM/I data for rain/no-rain pixels classification, and then GOES data to calibrate the relationship between brightness temperature and rain rate via linear regression for the precipitating pixels [39,40]. The PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) algorithm of the Center for Hydrometeorology and Remote Sensing (CHRS) is an adaptive, multi-platform precipitation estimation algorithm, based on an artificial neural network approach. It merges high quality data from National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), and Defense Meteorological Satellite Program (DMSP) low-altitude polar-orbit satellites with sampled data from geosynchronous satellites [41–43]. The Integrated Multi-satellitE Retrievals for GPM (IMERG) is a merged precipitation product developed by the US GPM science team. This algorithm is intended to produce fine time- and space-scale estimates for the entire globe using inter-calibrated, merged, and interpolated data from all available PMW satellites, together with microwave-calibrated infrared (IR) satellite estimates, precipitation gauge analyses, and other precipitation estimators [44].

The combination of MW and IR measurements generally follows two main techniques—the so-called "blended" or "microwave-calibrated" and "morphing". The first one is based on a calibration of IR cloud top temperatures measurements using the MW (namely Passive MW-PMW) precipitation estimates, in order to generate local relationships between the IR and PMW observations [31,32,35,43,45–50]). The derived relationships are then applied to the IR data, increasing the spatial and temporal extent of the precipitation estimation with respect to the PMW overpasses. The "morphing" technique is based on the evidence that IR data, locally updated using PMW-based rainfall measurements, can be employed to measure cloud movement, propagating forward in time the rain field, between the consecutive LEO PMW satellite overpasses [37,51–54]. Basically, this technique derives estimates of precipitation from infrared data when passive microwave information is unavailable.

This paper describes an algorithm, named RAINBOW (RAdar INfrared Blending algorithm for Operational Weather monitoring) combining the data collected by SEVIRI and by the Italian ground-based radars network, coordinated by the Italian Department of Civil Protection (IT GR) to provide precipitation estimation over Italy. The main objective of the algorithm is to provide rainfall estimates from SEVIRI observations, by exploiting the portion of IT GR data with the highest quality. The algorithm has been developed by using the "blended" approach taking using the Surface Rainfall Intensity (SRI) composite product obtained by combining the measurements from all the radars of the network. The Italian ground radar network represents a valuable monitoring system for the detection and warning of severe weather and related hydro-geological risks. As a matter of fact, Italy, and more generally the Mediterranean basin, is affected by severe weather events of different nature (e.g., deep convective systems, cyclones, tropical-like cyclones, etc.) hitting coastal as well as inland areas, causing serious damages and casualties [55–62]).

The IT GR is also currently an important part of the ground reference system for the Precipitation Product Validation Group of the EUMETSAT Satellite Application Facility for Support to Operational Hydrology and Water Management project [63]. However, the spatial heterogeneity of the data quality, related to orography and spatial coverage of the IT GR network, imposes the selection of the data to be used for blending.

The RAINBOW algorithm presented in this paper has been developed within the agreement between the Italian Department of Civil Protection and the Institute of Atmospheric Sciences and Climate (ISAC) of the National Research Council of Italy (CNR). The concept is to design an operational product to complement the radar monitoring of relevant precipitation events by covering both sea areas (not covered by IT GR) and areas where the quality of IT GR data is lower due to limited coverage and orographic obstruction. One of the request that has to be satisfied by RAINBOW is the as short as possible running time in order to provide precipitation estimates as soon as the SEVIRI acquisition is available.

This paper is organized as follows. Section 2 presents the instrumentation and methodology used in the design of the algorithm. Section 3 reports the results obtained by the algorithm when it is applied to selected case studies with the relative discussion. The conclusions are then reported in Section 4.

2. Instrumentation and Methods

Two-and-a-half years of data (from 1 July 2015 to 31 December 2017) collected by the IT GR network, and by the SEVIRI radiometer, have been used to develop the RAINBOW algorithm. The algorithm combines the SEVIRI brightness temperature and the precipitation rate estimated from the ground radars (GRs) to derive a relationship between these two quantities to be applied to each SEVIRI acquisition (i.e., every fifteen minutes). The area of interest is centered on the Italian peninsula, namely between 36–48°N and 6–20°E.

2.1. IT GR Network

At the time of the work, the Italian ground radar (GR) network includes 20 C-band and 3 X-band radar, managed by 11 administrations. Moreover, 7 C-band and 3 X-band systems (all with dual-polarization capability) are managed by the Department of Civil Protection (DPC), which is also the developer and distributor of the national precipitation product. The spatial distribution of the IT GR network with the associated Quality Index (QI) is depicted on Figure 1. The processing architecture is basically composed of two main steps, where the radar measurements are first locally processed by a unique software system, then all the products are centralized to generate the national level products.

There are different sources that can increase the uncertainty in the radar precipitation estimation [64]. The main errors can be identified by contamination by non-weather returns (clutter), partial beam blocking, beam broadening at increasing distances, vertical variability of precipitation [32,65,66], and rain path attenuation [1,67–69]. Due to the morphology of the Italian territory, the uncertainty can be mainly associated to the orography-related effects, especially in southern Italy where the radar coverage as well as the radar overlapping is poor [3,70]. Another error source is the Radio Local Area Network (RLAN) interferences, which are properly dealt with and filtered out using an effective algorithm based on a multi-parameter fuzzy logic approach that also make use of the Signal Quality Index (SQI).
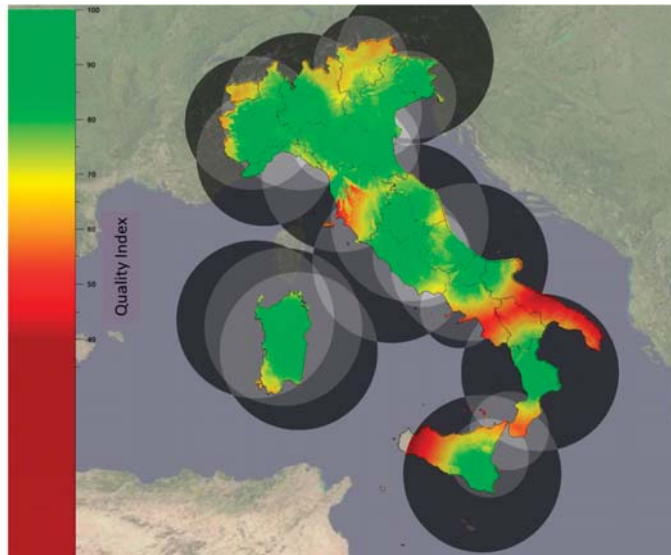
Figure 1. Italian ground radar network (IT GR) spatial coverage and its associated Quality Index (QI).

The processing system aims at identifying most of the uncertainty sources in order to compensate them, whenever it is possible, before estimating precipitation. As described in [71], the data quality index results from the combination of the partial QIs associated to each identified error source. A point-by-point description of the operational radar processing chain can be found in [72]. A sensitivity analysis, previously conducted, compared hourly rain gauge and radar data for increasing QI values. The results evidenced that the error (i.e., the difference between radar and rain gauges estimates) has its minimum value for QI = 0.60. At higher QI values, the error increased because of the presence of outliers together to a marked decrease of the sample size [72]. Following this analysis, only QI values equal or greater than 0.60 are considered reliable and are used within RAINBOW algorithm.

Furthermore, a filtering process is applied to the GR data by comparing them with the data collected by the Italian rain gauges network. The radar data showing marked differences with respect to the rain gauges measurements are discarded. The analysis is based on the ratio between the rain gauges and ground radars hourly cumulated data. Namely, the GR data are discarded if the ratio is less than 0.1, being this value chosen because it is much smaller than the average value that the ratio assumes close to the location of a calibrated polarimetric weather radar [73]. At the end of this operational process chain, the Surface Rainfall Intensity (SRI) product is provided over a 1×1 km^2 grid with a temporal resolution of 10 min. The SRI is obtained taking into account the orography (and the clutter associated), the technical characteristics of the radar (e.g., the various elevation angles and the scanning time frequency, the correction of the partial beam blocking [74,75]. In particular, the single-site SRI is estimated considered the whole radar volume in polar coordinates, then the national composite is computed in Cartesian coordinates. For a given geographical location, the single site SRI is retrieved combining the radar observations at all elevation scans θ_k, through a quality-weighted average [71,75,76]. Finally, the national SRI composite is built by combining the single-radar rainfall maps through a quality-weighted approach. In case in a given geographical location two or more radar SRI estimates are available, the one with the highest quality weights more.

2.2. SEVIRI Radiometer

The Spinning Enhanced Visible and InfraRed Imager (SEVIRI) radiometer [77] is the main instrument onboard of Meteosat Second Generation (MSG). The MSG is a geostationary satellite located at about 36,000 km above the Earth surface at 0°N, 0°E. SEVIRI is a passive microwave instrument collecting radiation from a target area and focusing it on detectors sensitive to 12 different bands of the electromagnetic spectrum. The twelve channels are distributed among visible part of electromagnetic spectrum (channels VIS 0.6 μm and VIS 0.8 μm), near-infrared (channel NIR 1.6 μm), infrared (channels IR 3.9 to IR 13.4 μm—for a total of eight channels) and High Resolution Visible (channel HRV 0.75 μm). The SEVIRI nominal time resolution is 15 min, of which twelve minutes are allocated to collect images, while the remaining three minutes are used for calibration, retrace, and stabilization. The SEVIRI spatial resolutions ranges from 1 km for the HRV channel to 3 km for VIS-NIR-IR channels at sub-satellite point (i.e., at 0°N, 0°E). The spatial resolution decreases moving away from the sub-satellite point (e.g., over the study area, the Italian peninsula, it is around 4 km for the VIS-NIR-IR channels).

SEVIRI measures the radiation emitted by a target located along the radiometer field of view (i.e., the total radiation emitted by clouds). Depending on the considered channel, the amount of measured radiation is representative of different cloud characteristics. While the measurement in the VIS channels gives an indication about the optical depth of the cloud, the measurement in the IR channels are generally indicative of different cloud properties. In this study, we focused only on three IR channels, namely channels 5, 6, and 9. Channels 5 and 6 are centered in the emission spectrum of the water vapor (WV) at wavelengths at 6.25 and 7.35 μm, respectively, giving an indication about the cloud optical depth other than to determine the water vapor distribution in two distinct layers of the atmosphere. The IR 10.8 μm channel provides continuous observation of the cloud top temperature. For these channels, the final output of SEVIRI is the brightness temperature (TB) that is defined as the temperature of a black body, which emits the same amount of radiation as observed.

2.3. P-IN-SEVIRI

P-IN-SEVIRI is a precipitation product developed in the Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF) project [78], providing instantaneous precipitation rate at spatial and temporal SEVIRI resolution. It is provided by EUMETSAT, and it is based on an underlying collection of time and space overlapping overpasses from SEVIRI IR imagers and surface rain rate estimates (through the use of algorithms based on Low Earth Orbit-Passive Microwave (LEO PMW) radiometers), which constitutes a look up table of geo-located relationships between rain rate and TB at 10.8 μm, updated as soon as new overlapping SEVIRI IR and LEO PMW overpasses are available. The processing method is called "Rapid Update" (RU) blending technique [79].

As new input datasets (MW and IR) are available in the processing chain, the MW-derived rain rat (RR) pixels are paired with their time and space-coincident geostationary 10.8 μm IR TB data, using a 10-min maximum allowed time offset between the pixel acquisition times and a maximum space offset of 10 km between the pixel coordinates. Each co-located data increments the histograms of TB and RR within a latitude-longitude box 2.5° wide (i.e., a 2.5° × 2.5° box), as well as the eight surrounding boxes (this overlap ensures a fairly smooth transition in the histogram shape between neighboring boxes). The rationale behind these threshold values for time collocation and box size is discussed by [80].

In order to set-up a meaningful statistical ensemble, the method can look at older MW-IR slot intersections (no older than 24 h), until a certain (75%) box coverage is reached and a minimum number of coincident observations are gathered for a 2.5° × 2.5° region (at present 400 points, this is a tunable parameter in the procedure). Thus, the RU technique requires an initial start-up time period (~24 h), to allow for establishing meaningful, initial relationships all over the considered area.

As soon as a box is refreshed with new data, a probabilistic histogram matching relationship is updated using the MW RR and IR TB probability distribution functions (PDF), and an updated lookup table (histogram file) is created.

2.4. GRISO

The Random Generator of Spatial Interpolation from uncertain Observations (GRISO) [81,82] is an improved kriging-like technique implemented by the International Centre on Environmental Monitoring (CIMA Research Foundation) to provide rainfall rate estimates. As input, GRISO uses the data from the Italian rain gauge network composed by roughly 3000 tipping bucket gauges (the number can change because of new instrument installation or malfunctioning of the available ones). While, in general, the rain gauge temporal sampling can change, instrument-by-instrument, ranging between 1 to 60 min (the minimum sampling time for Italian rain gauges is set to 15 min), the minimum detectable rain amount is equal to 0.2 mm. The GRISO technique preserves the rainfall rate values measured at the gauge location, allowing for a dynamical definition of the covariance structure associated with each rain gauge by the interpolation procedure. Each correlation structure depends both on the rain gauge location and on the accumulation time considered. Furthermore, GRISO is adopted in the H SAF validation procedure in comparison with European ground data [63] and respect to Dual-frequency Precipitation Radar (DPR) precipitation product [72]. The GRISO data available are provided over a regular grid (1 km × 1 km) with an hourly time step.

2.5. Parallax Correction

As highlighted in Sections 2.1 and 2.2, IT GR has higher spatial resolution than SEVIRI (i.e., 1 km versus to 4 km). The first step to correctly match ground-based radar and satellite observation is the upscale of the IT GR data to the SEVIRI resolution. Preliminarily, it has to be highlighted that satellite observations of the top surface of clouds is affected by the parallax effect (parallax error), which results in a dislocation of the ground mapped position. The parallax error is a function of three factors that is latitude, longitude, and height of the cloud other than the radius of Earth. While latitude and longitude of the cloud and radius of Earth are known, the height of the cloud has to be determined.

To this end, the TB measured by SEVIRI channel 9 ($TB_{10.8}$) is matched with the vertical profiles of temperature provided by European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-Interim) data [83–85]. The ERA-Interim data are provided on the same grid of SEVIRI over 37 not equi-spaced pressure levels (from 10 to 1000 hPa corresponding to altitudes ranging from 0 to 16 km about with spatial resolution between 240 and 1400 m about) with a time resolution of six hours (i.e., four runs of the model per day). For each SEVIRI instantaneous field of view (IFOV), the $TB_{10.8}$ is compared with the corresponding and closest in time vertical profile of temperature provided by ERA-Interim in order to estimate the cloud top height. At this point, the formula reported by Equation (1) can be applied to quantify the parallax displacement as function of longitude and latitude:

$$\Delta\gamma(\lambda, \phi) = \frac{P \cdot \sqrt{1 - \cos^2\lambda \cdot \cos^2\phi}}{P \cdot \cos\lambda \cdot \cos\phi - 1} \cdot \frac{h}{R} \tag{1}$$

where $P = 1 + \frac{H}{R}$ with H distance between satellite and Earth surface (~36,000 km), R radius of Earth, h height of cloud top, λ and ϕ longitude and latitude, respectively. Once that $\Delta\gamma(\lambda,\phi)$ is calculated, it can be converted in number of SEVIRI IFOV displacement both in longitude and latitude. The cloud is then moved to the correct position. The parallax displacement can be marked over the Mediterranean area depending on the cloud top height.

Figure 2 shows the parallax displacement (in km) as function of latitude, longitude and cloud top height. The parallax displacement for low clouds is almost constant around 2.3 km, regardless of the coordinates (latitude, longitude) of the measurement point. For higher cloud top, the displacement becomes significant (up to 15–20 km), depending also on the geographical position. The displacement varies by about 5/6 km for cloud heights of 11/14 km moving from south to north (i.e., from 36°N to 46°N and at a given longitude). Moving from west to east (and, therefore, at the same latitude), the variability of the parallax displacement is more limited (from about 1.5 to 2 km).
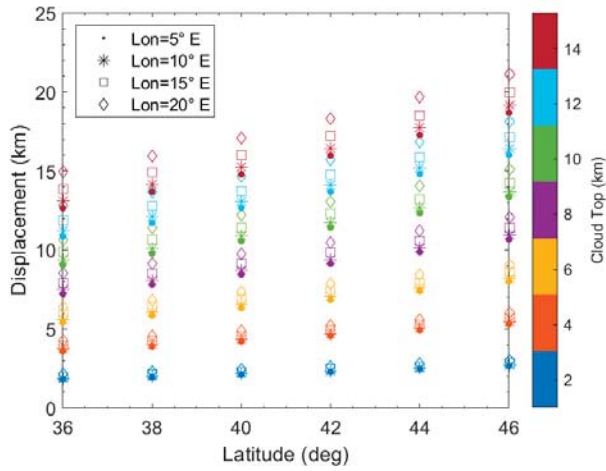
Figure 2. Parallax displacement as function of latitude, longitude, and cloud top height.

2.6. RAINBOW Algorithm

The RAINBOW algorithm is composed by a static module, which has been developed using historical data, and a dynamic module, which continuously updates the data to be used.

Both static and dynamic modules of RAINBOW have been developed for each of the five geographical boxes in which the area of interest has been divided (Figure 3). The choice to divide the area of study in geographical boxes is mainly related to the fact that precipitations with different microphysics properties can occurred over the Italian territory (e.g., a precipitation over the Alps may have different characteristics of a simultaneous precipitation over sea and/or in proximity of the coast). In addition, the precipitation occurring at the same time in different locations could be at different stage of its evolution. Dividing the area of study in geographical boxes mitigates the problems deriving from the situations just above described. In general, the smaller the box the better is the characterization of the precipitation. However, the box size has to be large enough to ensure an adequate number of samples in order to perform a reliable calibration. At the same time, an excessive number of geographical boxes can create discontinuities in the transition zones (i.e., on the line connecting two adjacent boxes). It was found that a good trade-off for the Italian country was to divide the country in five boxes.
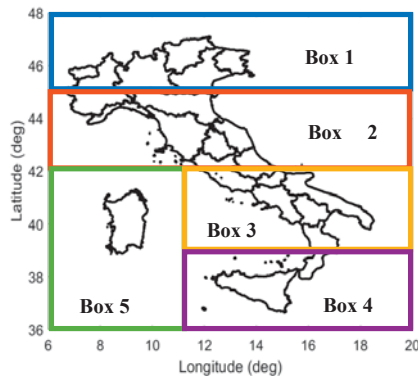
Figure 3. Geographical boxes division of the area of study.

The RAINBOW algorithm works with data at SEVIRI spatial and temporal resolution and provides the output at the same spatial and temporal resolution. Thus, the first step is to downscale the SRI data at SEVIRI resolution. The SRI pixels selected for each SEVIRI IFOV have to satisfy two different thresholds:

The mean QI is calculated considering all the IT GR pixels within a SEVIRI IFOV. To consider the IFOV useful, the mean QI has to be higher than 60%.

If the threshold of 60% for the mean QI is overcome, the mean SRI (i.e., the mean precipitation rate for a SEVIRI IFOV) is calculated by considering only the pixel with QI $\geq$ 80%. The threshold at 80% allows to discard the pixels affected by any possible spurious signal (e.g., noise, beam blockage, etc.). At the same time, the maximum SRI value is stored.

At this point, RAINBOW decides if to use the static or dynamic part of the algorithm. The decision is based on the number of useful IFOVs in each geographical box (i.e., the IFOVs with both RR and TB data) collected both in the last hour with respect to the running time and in the last SEVIRI acquisition (we recall that GR data have higher temporal resolution than SEVIRI, ten versus 15 min, respectively). In particular, if the number of useful IFOVs in the last hour is higher (or equal) than 50% and, the number of useful IFOVs in the last acquisition is higher (or equal) than 10% or lower than 10% but the maximum RR exceed 3 mmh^{-1}, the dynamic module of RAINBOW algorithm is applied. On the other hand, if these conditions are not satisfied, the static module of RAINBOW algorithm is used. The thresholds are defined through sensitivity tests changing both the percentage of useful IFOVs and the maximum RR value. The final output of both dynamic and static part of RAINBOW is a RR-TB$_{10.8}$ relationship, for each geographical box, to be applied to the SEVIRI data in order to give precipitation estimation. The main difference between the two modules is that the dynamic one updates and changes the RR-TB$_{10.8}$ relationship at each new SEVIRI acquisition, while the static one makes use of RR-TB$_{10.8}$ relationships obtained by considering the whole dataset available (i.e., from 1 July, 2015 to 31 December, 2017). Furthermore, a RR-TB$_{10.8}$ relationship for each meteorological season is derived in the static module. The RR-TB$_{10.8}$ relationship is obtained by sampling the TB$_{10.8}$ between 200 K and 270 K in 35 bins 2 K width. For each bin, the mean rainfall rate and the mean of maxima rainfall rates are calculated. More specifically, the TB$_{10.8}$ spectrum is split in two parts, one between 200 K and 220 K and one between 220 K and 270 K, and two RR-TB$_{10.8}$ relationships are derived. A second degree polynomial RR-TB$_{10.8}$ relationship is derived for the first part of TB$_{10.8}$ spectrum ($200 \leq$ TB$_{10.8}$ ≤ 220 K), while a first degree polynomial RR-TB$_{10.8}$ relationship is derived for the first part of TB$_{10.8}$ spectrum ($220 <$ TB$_{10.8}$ ≤ 270 K).

Figure 4 shows, as an example, the RR-TB$_{10.8}$ relationship obtained from the whole dataset for each season and each box used by the static module of the algorithm. It outlines how the higher rainfall rates are associated to the lower TB$_{10.8}$. Fall and summer (Figure 4a–d) are the seasons where this relationship is more straightforward for all the considered geographical boxes. At the same time, winter (Figure 4b) is the season with the lowest precipitation rate (as could be expected) and with a very light relationship between RR and TB$_{10.8}$. Together to the RR-TB$_{10.8}$ relationship, the probability of precipitation (POP) is calculated for each TB$_{10.8}$ bin and the corresponding POP-TB$_{10.8}$ relationship is derived. The POP is defined as the ratio between the number of SEVIRI IFOVs with precipitation (RR ≥ 0.25 mmh^{-1}) and the number of SEVIRI IFOVs with no precipitation (RR < 0.25 mmh^{-1}). As for the RR-TB$_{10.8}$ relationship, the dynamic module of RAINBOW updates and changes the POP-TB$_{10.8}$ relationship at each SEVIRI acquisition, while the static module again takes advantages of the POP-TB$_{10.8}$ relationship (for each box and each season) built by using the whole available dataset.
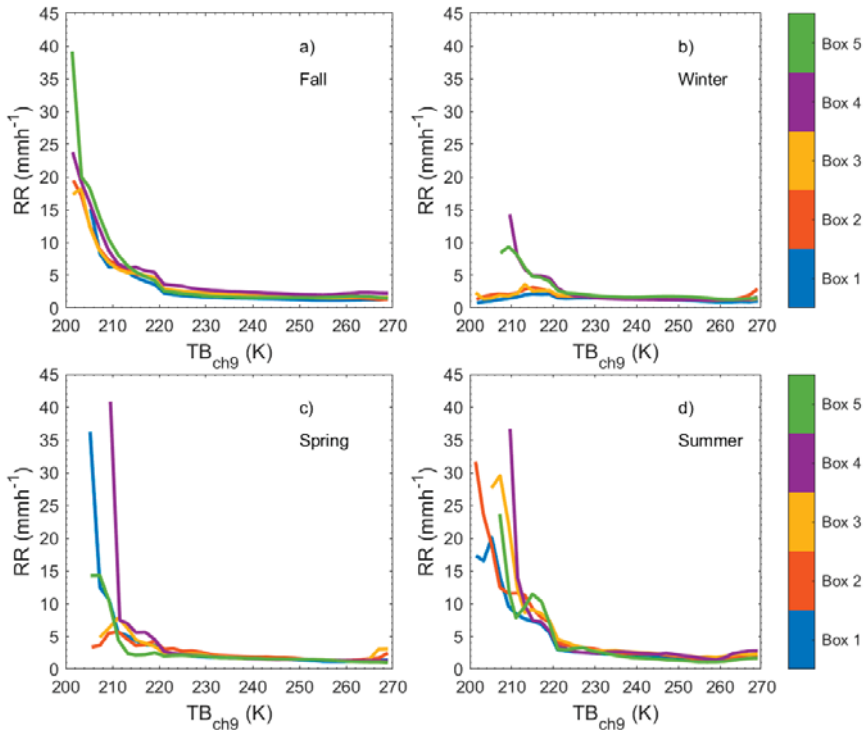
Figure 4. RR-TB$_{10.8}$ relationship obtained from the whole dataset (i.e., from 1 July, 2015 to 31 December, 31, 2017) for each box for (**a**) fall, (**b**) winter, (**c**) spring, and (**d**) summer season, respectively.

Figure 5 reports the POP-TB$_{10.8}$ relationships derived for each season and each box. The POP clearly increases decreasing the TB$_{10.8}$ during fall and summer season (Figure 5a–d), reaching the 100% for TB$_{10.8}$ as low as 210 K (boxes 2 and 3 show a decrease of POP for TB$_{10.8}$ < 210 K during fall season—Figure 5a). Not as straightforward as for fall/summer is the POP-TB$_{10.8}$ relationship for spring/winter (Figure 5b–c). There is a sharp decrease of POP at TB$_{10.8}$ higher than 255 K. At the same time, POP increases decreasing TB$_{10.8}$ up to 220 K about; then, the trend diversifies among the boxes, with most of them showing a marked decrease of POP for TB$_{10.8}$ lower than 220 K. Among these, someone present a sharp increase when TB$_{10.8}$ reaches values lower than 210K. The decrease of POP at lower TB$_{10.8}$ values is mainly related to the presence of cirrus clouds, which are no-precipitating clouds with very low cloud top temperature. The occurrence of cirrus clouds reaches a maximum (minimum) in winter (summer) [86]. This aspect is related to the lower temperature in the troposphere during winter that favors both the formation and the maintenance of ice crystals, which are the constituents of this type of clouds [87].
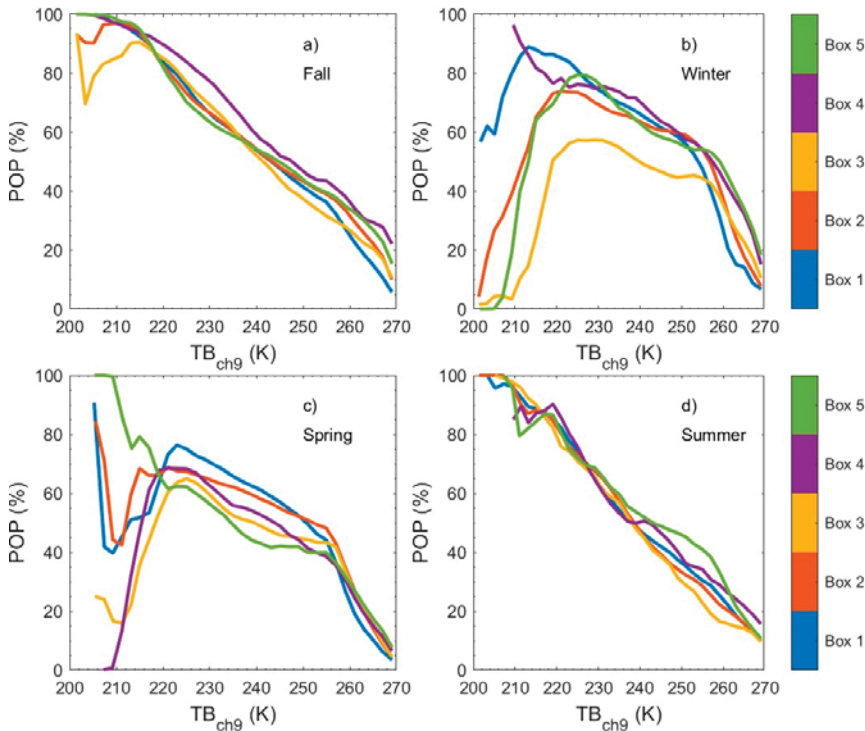
Figure 5. POP-TB$_{10.8}$ relationship obtained from the whole dataset (i.e., from 1 July, 2015 to 31 December, 2017) for each box for (**a**) fall, (**b**) winter, (**c**) spring, and (**d**) summer season, respectively.

3. Results

The methodology described above has been applied to several case studies. The algorithm performances were analyzed by comparing the RAINBOW precipitation retrievals with the outputs of GRISO and P-IN-SEVIRI on a regular grid (0.25° × 0.25°) for ten selected case studies (occurred in 2016 and 2017). Furthermore, the potentialities and limitations of RAINBOW are discussed for two outputs of the algorithm considering two different case studies.

The first considered event occurred in the night between 9 and 10 September, 2017, causing a flash flood which hit the coastal city of Livorno (43.5°N, 10.3°E), in the Tuscany region. In the area around the city, three rain gauges measured more than 230 mm of accumulated precipitation in six hours (00:00–06:00 UTC), with peaks of 150 mm h^{-1} registered between 01:00 and 03:00 UTC.

Regarding the event observed on 10 September 2017, Figure 6 shows the TB$_{10.8}$ as measured by SEVIRI (Figure 6a), the instantaneous rainfall rate as estimated by IT GR network at SEVIRI spatial resolution (Figure 6b) and by RAINBOW (Figure 6c) at 01:12 UTC, respectively. The SEVIRI TB$_{10.8}$ (Figure 6a) highlights the presence of a V-shaped thunderstorm hitting mainly the north part of Tuscany region. The updraft core developed over sea, just offshore of the coastal line remained stationary for several hours (roughly between 18:00 UTC of 9 September and the 03:00 UTC of 10 September). Values of TB$_{10.8}$ as low as about 210 K are measured in the updraft core corresponding to a cloud top height around 12 km. The plot also outlines the presence of a storm line across the Sardinia region. The IT GR network estimated rainfall rate values up to 50 mmh^{-1} (Figure 6b) within a SEVIRI IFOV (i.e., round 4 km × 4 km). At the same time, the spatial extension of the storm is quite limited both in terms of cloud and precipitation coverage. The same can be said for the precipitation across

the Sardinia even if the estimated rainfall rates reach lower values up to 40 mmh^{-1}. Finally, lighter precipitation is detected in the northern part of Italy. The RAINBOW rainfall rate estimation (Figure 6c) captures well the two most intense precipitation zones (i.e., the area around Livorno and over Sardinia) but tends to detect precipitation over a larger area than radar. At the same time, the precipitation peak is well identified in both location and intensity, with a slight underestimation of the most intense cells.
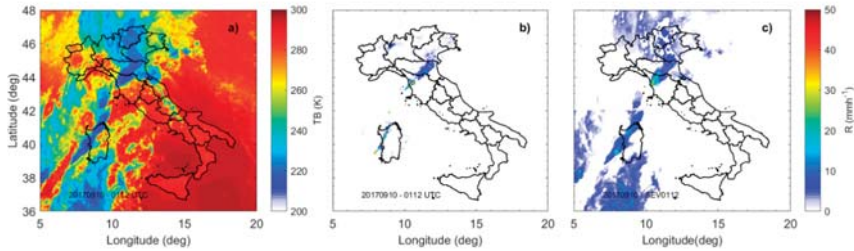
Figure 6. Snapshot relative to the 01:12 UTC of 10 September, 2017. Panel (**a**) shows the TB10.8 as measured by Spinning Enhanced Visible and InfraRed Imager (SEVIRI), (**b**) the instantaneous rainfall rate as estimated by Italian ground radar network (IT GR) network at SEVIRI spatial resolution and (**c**) by RAdar INfrared Blending algorithm for Operational Weather monitoring (RAINBOW).

Figure 7 shows a snapshot relative to the 04:12 UTC for the case study of 14 October, 2016. Although the storm involved the same region (at least at that time), different properties of RAINBOW can be highlighted by the analysis if this case study. The case study reported in Figure 7 presents different characteristics showing two convective cells, one between Tuscany and Emilia Romagna regions, and one out of the Italian territory over south France (partially over sea and partially over land). Both convective cells have bigger spatial extension and even colder TB$_{10.8}$ values up to 205 K about (Figure 7a). To the big cloud extension does not correspond an equal precipitation extension; in fact, the IT GR network shows scattered and small precipitation clusters with a quite wide range of intensity from few mmh^{-1} to almost 50 mmh^{-1} (Figure 7b). Analyzing the precipitation estimated by RAINBOW, it is possible to note significant differences with respect to SRI (Figure 7c):
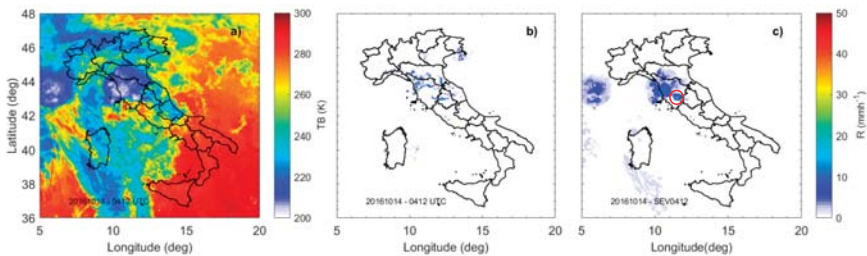
Figure 7. Snapshot relative to the 04:12 UTC of 14 October, 2016. Panel (**a**) shows the TB$_{10.8}$ as measured by SEVIRI, (**b**) the instantaneous rainfall rate as estimated by IT GR network at SEVIRI spatial resolution and (**c**) by RAINBOW.

The rainfall rate peak estimated by RAINBOW is weaker than that estimated by IT GR, with maximum values around 20 mmh^{-1}. This can be mainly attributed to the limited number of IFOVs with intense rainfall rate considered in the calibration process.

RAINBOW is able to estimate precipitation for the convective cell over France and for the small cell on the border between Tuscany and Umbria region (red circle in Figure 7c). However, the precipitation corresponding to this latter cell is slightly overestimated, in terms of spatial extension, by RAINBOW.

On the other hand, the precipitation cluster centered on the coastal line of Tuscany is well detected by RAINBOW. In the operational frame in which the algorithm is intended, this case study highlights the potentialities of RAINBOW. The precipitation detection of the two cells can be considered as warning of a possible event moving toward the Italian territory and as complementary to the SRI estimation, respectively.

4. Discussion

The algorithm performances were assessed by comparing the RAINBOW outputs with the GRISO data (taken as reference) on a regular 0.25° × 0.25° grid for 10 case studies. Since both RAINBOW and GRISO are provided at higher but different spatial resolutions, they are up-scaled to a regular 0.25° × 0.25° grid. Both categorical scores (Probability of Detection (POD), False Alarm Ratio (FAR), Heidke Skill Score (HSS)) and continuous scores (mean error (ME) and root mean square error (RMSE)) have been considered [88]. The analysis has been done on an hourly basis (mm of rain fell in this time interval) considering the entire event of each case study. Furthermore, a minimum cumulative hourly rainfall threshold of 0.25 mm and three different intervals of cumulated rain are considered: light 0.25–1 mm, moderate 1–10 mm, and heavy 10–100 mm. The statistical scores above reported have been calculated even between P-IN-SEVIRI and GRISO in order to compare the RAINBOW and P-IN-SEVIRI performances.

The results shown in Figure 8 evidence excellent algorithm performance especially for moderate and heavy precipitation intensity. The Probability of Detection (POD)—Figure 8a) ranges between 0.8 and 1, except for light precipitation (0.25–1 mm); the False Alarm Ratio (FAR)—Figure 8b) has a specular trend with respect to the POD, with higher values for light precipitation and lower for the other rain intervals, while the Heidke Skill Score (HSS)—Figure 8c) follows the trend of the POD with values up to 0.8. It should be noted that the values of POD, FAR, and HSS are almost constant for all 10 case studies, underlining an excellent stability of the algorithm. In particular, HSS increases with time, highlighting that the continuous update of DPR GR network plays a crucial role in the RAINBOW performance by supplying ever-higher quality data input.
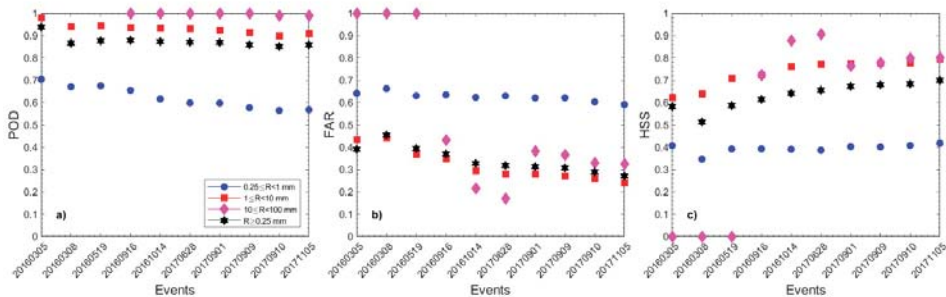
Figure 8. (**a**) Probability of Detection (POD), (**b**) False Alarm Ratio (FAR), and (**c**) Heidke Skill Score (HSS) scores calculated by comparing the RAINBOW outputs with the Random Generator of Spatial Interpolation from uncertain Observations (GRISO) data (taken as reference) on a regular 0.25° × 0.25° grid for 10 case studies. A minimum cumulated rain threshold is set at 0.25 mm and three different intervals of cumulated rain are considered: light 0.25–1 mm, moderate 1–10 mm, and heavy 10–100 mm.

The algorithm error in estimating the precipitation rate is quantified with respect to GRISO by calculating the mean error (ME) and the root mean square error (RMSE). Figure 9a shows that the ME oscillates around 0 mm for all cases and for all precipitation intervals except for heavy intensity where the values range between −7 and −9 mm indicating a clear underestimation of the higher intensities by

the algorithm. The good results are confirmed by the RMSE (Figure 9b), which never exceeds 3 mm except for intense rainfall.
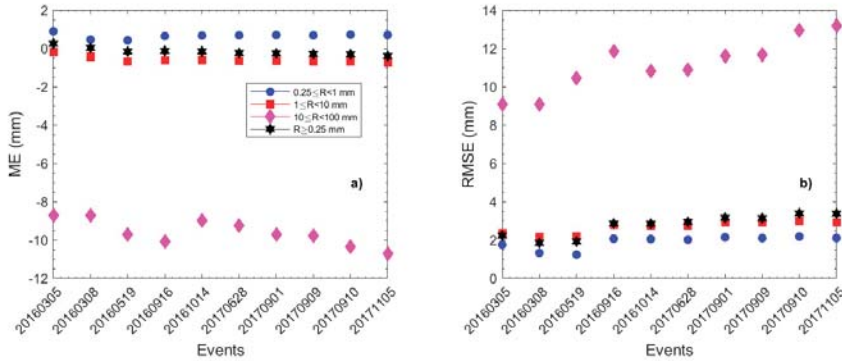
Figure 9. (**a**) Mean error (ME) and (**b**) root mean square error (RMSE) scores calculated by comparing the RAINBOW outputs with the GRISO data (taken as reference) on a regular 0.25° × 0.25° grid for 10 case studies. A minimum cumulated rain threshold is set at 0.25 mm and three different intervals of cumulated rain are considered: light 0.25–1 mm, moderate 1–10 mm, and heavy 10–100 mm.

A sensitivity study has been conducted in order to evaluate the performance of RAINBOW as a function of the resolution of the regular grid. To this end, four different grids have been chosen ranging from 0.1° × 0.1° to 0.25° × 0.25°. The analysis has been always done on an hourly basis considering only the minimum cumulative hourly rainfall threshold of 0.25 mm.

The results shown in Figure 10 evidence very stable values for the categorical scores as a function of the resolution of the grid. In particular, POD (Figure 10a) has constant values slightly higher than 0.8, while both FAR and HSS (Figure 10b,c, respectively) show a more irregular trend only for the 0.1° × 0.1° grid with higher and lower values, respectively, than the other grids.
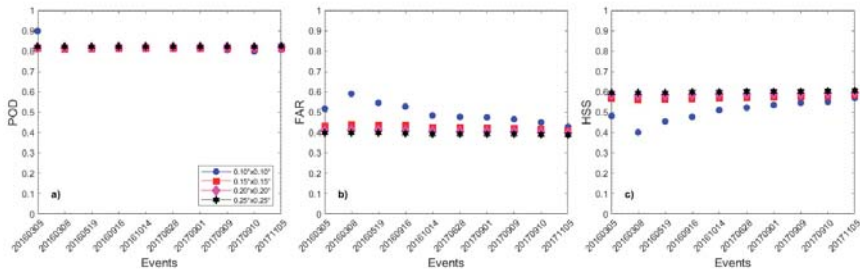
Figure 10. (**a**) POD, (**b**) FAR, and (**c**) HSS scores calculated by comparing the RAINBOW outputs with the GRISO data (taken as reference) on different regular grids for 10 case studies. A minimum cumulated rain threshold is set at 0.25 mm.

The continuous scores in Figure 11 confirm the results shown in Figure 10. The ME (Figure 11a) is always negative, around −0.4 mm, except for the first two case studies of 0.1° × 0.1° grid. On the other hand, the RMSE (Figure 11b) has very limited variations around 3.2 (mm), while for 0.1° × 0.1° grid, it shows an irregular trend with values dropping down up to 1.8 mm.
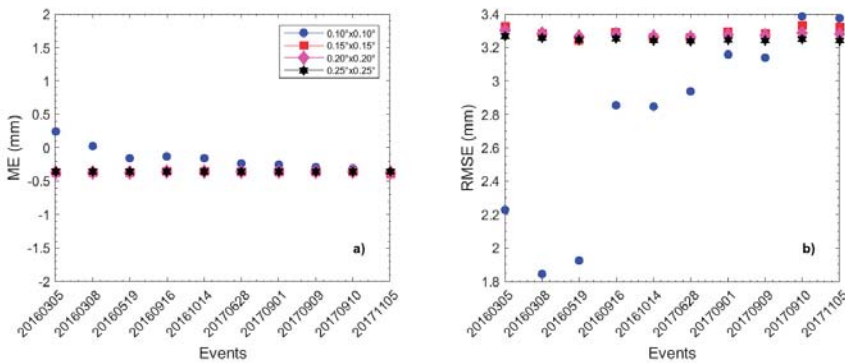
Figure 11. (**a**) ME and (**b**) RMSE scores calculated by comparing the RAINBOW outputs with the GRISO data (taken as reference) on different regular grids for 10 case studies. A minimum cumulated rain threshold is set at 0.25 mm.

The same analyses, shown in Figures 8 and 9, have been carried out by comparing the statistical scores calculated for RAINBOW with those one calculated for P-IN-SEVIRI product (always taking GRISO as reference). The results are ported in Figures 10 and 11 for categorical and continuous scores, respectively.

Figure 12 evidences the better performances of RAINBOW in detecting precipitation. The POD$_{RAINBOW}$ is always higher than POD$_{P-IN-SEVIRI}$ (Figure 12a) regardless the intensity of precipitation (different marker shape in the plot) and the different events (labeled by different colors). For light precipitation (circle markers), the POD$_{P-IN-SEVIRI}$ does not exceed 0.3, while POD$_{RAINBOW}$ ranges between 0.5 and 0.7. At moderate and heavy precipitation (and even not considering any rain intervals), while POD$_{RAINBOW}$ is always above 0.8, POD$_{P-IN-SEVIRI}$ shows a wide range of values between 0.2 and 1. At the same time, the FAR is very similar between the two algorithms with most of the points on the one-to-one line and at values generally lower than 0.4 (Figure 12b). The combination of POD and FAR results in constantly higher values oh HSS$_{RAINBOW}$ with respect to P-IN-SEVIRI (Figure 12c). The very good performances of RAINBOW in detecting the precipitation are confirmed by continuous scores, which refer to the precipitation rate estimation.
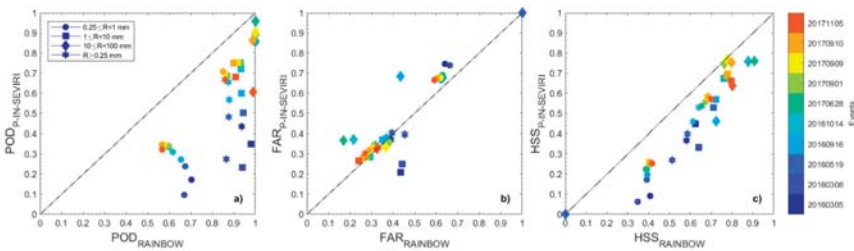
Figure 12. Comparison of (**a**) POD, (**b**) FAR, and (**c**) HSS scores calculated for RAINBOW and P-IN-SEVIRI outputs with respect to the GRISO data (taken as reference) on a regular 0.25° × 0.25° grid for 10 case studies. A minimum cumulated rain threshold is set at 0.25 mm and three different intervals of cumulated rain are considered: light 0.25–1 mm, moderate 1–10 mm, and heavy 10–100 mm.

Figure 13a shows that ME$_{RAINBOW}$ and ME$_{P-IN-SEVIRI}$ are very similar for heavier precipitation intensity, while ME$_{RAINBOW}$ and ME$_{P-IN-SEVIRI}$ assume values around 0 mm and slightly negative, respectively, for light to moderate precipitation intensity. On the other hand, RMSE$_{RAINBOW}$ is generally lower than RMSE$_{P-IN-SEVIRI}$ regardless the precipitation rate (Figure 13b).
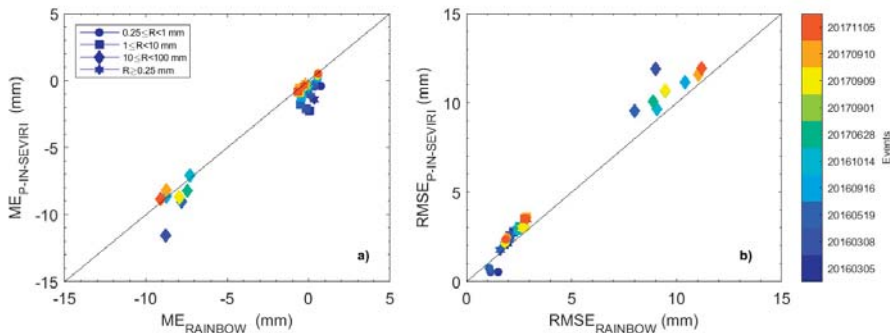
Figure 13. Comparison of (**a**) ME and (**b**) RMSE scores calculated for RAINBOW and P-IN-SEVIRI outputs with respect to the GRISO data (taken as reference) on a regular 0.25° × 0.25° grid for 10 case studies. A minimum cumulated rain threshold is set at 0.25 mm and three different intervals of cumulated rain are considered: light 0.25–1 mm, moderate 1–10 mm, and heavy 10–100 mm.

5. Conclusions

A new algorithm (RAINBOW) based on the combination of the data collected by SEVIRI onboard of MSG) and by the Italian ground-based radars network (IT GR) to provide precipitation estimation over Italy has been described. The algorithm, consisting of two main modules and operating over five geographical boxes in which the study area is divided, derives and updates (whenever it is possible) second degree polynomial RR-TB$_{10.8}$ relationships. These relationships are applied to each acquisition of SEVIRI in order to provide a precipitation map. The results, based on a number of case studies, show good performance of the algorithm when it is compared with ground reference (i.e., GRISO precipitation pattern and intensity derived from rain gauge measurements), with high/low values for POD/FAR especially for light to moderate precipitation range. At the same time, the ME values are close to 0 mmh^{-1}, while RMSE is about 2 mmh^{-1}, highlighting a remarkable accuracy of RAINBOW estimates, whereas the capability to detect the precipitation pattern and intensity decreases for severe phenomena. It has to be remarked that severe events could be characterized by high spatial variability, which cannot be accomplished by RAINBOW (due to the SEVIRI instrument characteristics). It is worth noting that the performance of RAINBOW are quite constant through the different case studies with a slight improvement of the performance over time. This is related to the fact that RAINBOW relies on the high quality precipitation rate estimates from IT GR network, which are constantly maintained and upgraded. Furthermore, RAINBOW shows better performance than P-IN-SEVIRI (i.e., the H SAF product based on IR-derived precipitation estimation) when both products are compared to GRISO.

RAINBOW was conceived as an operational product to supply data where the IT GR coverage is absent or it presents low QI values. In this regard, the main aim of RAINBOW is the detection of extreme events that are barely observed by IT GR network in order to support the pre-alarm system for the hydro-geological risks and the life threatening conditions related to the incoming extreme events. Furthermore, the algorithm has to comply with short running time and with ease of management, which are fundamental aspects in a pre-alarm system. RAINBOW ensures running time comparable (or even shorter) with the IT GR running time and significantly shorter than P-IN-SEVIRI running time.

The next launch (scheduled in December 2021) of Flexible Combined Imager (FCI) on board of Meteosat Third Generation (MTG), will be useful to further improve the performance of RAINBOW. The higher number of channels available, the higher spatial and temporal resolution will provide higher quality data to characterize, also, very local severe events.

Author Contributions: Conceptualization, L.P.D. and S.D.; Methodology, L.P.D., P.S., and S.D.; Validation, L.P.D.; Formal Analysis, L.P.D.; Investigation, L.P.D.; Resources, S.D., S.P., G.V., and M.P.; Data Curation, L.P.D. and M.P.;

References

1. Bringi, V.N.; Chandrasekar, V. *Polarimetric Doppler Weather Radar: Principles and Applications*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2001; ISBN 978-0-521-62384-1.

2. Kirstetter, P.-E.; Delrieu, G.; Boudevillain, B.; Obled, C. Toward an error model for radar quantitative precipitation estimation in the Cévennes–Vivarais region, France. *J. Hydrol.* **2010**, *394*, 28–41. [CrossRef]

3. Vulpiani, G.; Montopoli, M.; Passeri, L.D.; Gioia, A.G.; Giordano, P.; Marzano, F.S. On the use of Dual-Polarized C-Band radar for operational rainfall retrieval in mountainous areas. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 405–425. [CrossRef]

4. Derin, Y.; Anagnostou, E.; Anagnostou, M.N.; Kalogiros, J.; Casella, D.; Marra, A.C.; Panegrossi, G.; Sano, P. Passive microwave rainfall error analysis using high-resolution X-Band Dual-Polarization radar observations in complex terrain. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2565–2586. [CrossRef]

5. Bennartz, R.; Thoss, A.; Dybbroe, A.; Michelson, D.B. Precipitation analysis using the Advanced Microwave Sounding Unit in support of nowcasting applications. *Meteorol. Appl.* **2002**, *9*, 177–189. [CrossRef]

6. Stephens, G.L.; Kummerow, C.D. The Remote Sensing of Clouds and Precipitation from Space: A Review. *J. Atmos. Sci.* **2007**, *64*, 3742–3765. [CrossRef]

7. Kidd, C.; Matsui, T.; Chern, J.; Mohr, K.; Kummerow, C.; Randel, D. Global precipitation estimates from cross-track passive microwave observations using a physically based retrieval scheme. *J. Hydrometeorol.* **2016**, *17*, 383–400. [CrossRef]

8. Levizzani, V.; Cattani, E. Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. *Remote Sens.* **2019**, *11*, 2301. [CrossRef]

9. Arkin, P.A.; Meisner, B.N. The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982-84. *Mon. Weather Rev.* **1987**, *115*, 51–74. [CrossRef]

10. Rickenbach, T.M. Cloud-top evolution of tropical oceanic squall lines from radar reflectivity and infrared satellite data. *Mon. Weather Rev.* **1999**, *127*, 2951–2976. [CrossRef]

11. Amorati, R.; Alberoni, P.P.; Levizzani, V.; Nanni, S. IR-based satellite and radar rainfall estimates of convective storms over northern Italy. *Meteorol. Appl.* **2000**, *7*, 1–18. [CrossRef]

12. Levizzani, V.; Schmetz, J.; Lutz, H.J.; Kerkmann, J.; Alberoni, P.P.; Cervino, M. Precipitation estimations from geostationary orbit and prospects for METEOSAT Second Generation. *Meteorol. Appl.* **2001**, *8*, 23–41. [CrossRef]

13. Lazri, M.; Ameur, S.; Brucker, J.M.; Testud, J.; Hamadache, B.; Hameg, S.; Ouallouche, F.; Mohia, Y. Identification of raining clouds using a method based on optical and microphysical cloud properties from Meteosat second generation daytime and nighttime data. *Appl. Water Sci.* **2013**, *3*, 1–11. [CrossRef]

14. Capacci, D.; Porcù, F. Evaluation of a satellite multispectral VIS–IR daytime statistical rain-rate classifier and comparison with passive microwave rainfall estimates. *J. Appl. Meteorol. Climatol.* **2009**, *48*, 284–300. [CrossRef]

15. Kühnlein, M.; Thies, B.; Nauß, T.; Bendix, J. Rainfall-rate assignment using MSG SEVIRI Data—A promising approach to spaceborne rainfall-rate retrieval for midlatitudes. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 1477–1495. [CrossRef]

16. Kühnlein, M.; Appelhans, T.; Thies, B.; Nauß, T. Precipitation estimates from MSG SEVIRI daytime, nighttime, and twilight data with random forests. *J. Appl. Meteorol. Climatol.* **2014**, *53*, 2457–2480. [CrossRef]

17. Feidas, H.; Giannakos, A. Identifying precipitating clouds in Greece using multispectral infrared Meteosat Second Generation satellite data. *Theor. Appl. Climatol.* **2011**, *104*, 25–42. [CrossRef]

18. Thies, B.; Nauß, T.; Bendix, J. Precipitation process and rainfall intensity differentiation using Meteosat Second Generation Spinning Enhanced Visible and Infrared Imager data. *J. Geophys. Res.* **2008**, *113*, D23. [CrossRef]

19. Anagnostou, E.N.; Kummerow, C. Stratiform and convective classification of rainfall using SSM/I 85-GHz brightness temperature observations. *J. Atmos. Ocean. Technol.* **1997**, *14*, 570–575. [CrossRef]

20. Tapiador, F.; Marcos, C.; Sancho, J. The convective rainfall rate from cloud physical properties algorithm for meteosat second-generation satellites: Microphysical basis and intercomparisons using an object-based method. *Remote Sens.* **2019**, *11*, 527. [CrossRef]

21. Jurczyk, A.; Szturc, J.; Otop, I.; Ośródka, K.; Struzik, P. Quality-based combination of multi-source precipitation data. *Remote Sens.* **2020**, *12*, 1709. [CrossRef]

22. NWC SAF. Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO. 2019. Available online: http://www.nwcsaf.org/Downloads/GEO/2018/Documents/Scientific_Docs/NWC-CDOP2-GEO-AEMET-SCI-ATBD-Precipitation_v2.1.pdf (accessed on 30 July 2020).

23. Kummerow, C.D.; Tanelli, S.; Takahashi, N.; Furukawa, K.; Klein, M.; Levizzani, V. Plans for Future Missions. In *Satellite Precipitation Measurement*; Advances in Global Change Research; Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 67, pp. 99–119. ISBN 978-3-030-24567-2.

24. Bernard, F.; Pasternak, F.; Davancens, R.; Baldit, E.; Luitot, C.; Penquer, A.; Calvel, B.; Buil, C. Overview of IASI-NG the new generation of infrared atmospheric sounder. In Proceedings of the International Conference on Space Optics—ICSO, Tenerife, Spain, 6–10 October 2014; Cugny, B., Sodnik, Z., Karafolas, N., Eds.; SPIE: Tenerife, Spain, 2017; p. 41.

25. Andrey-Andrés, J.; Fourrié, N.; Guidard, V.; Armante, R.; Brunel, P.; Crevoisier, C.; Tournier, B. A simulated observation database to assess the impact of the IASI-NG hyperspectral infrared sounder. *Atmos. Meas. Tech.* **2018**, *11*, 803–818. [CrossRef]

26. Mugnai, A.; Cooper, H.J.; Smith, E.A.; Tripoli, G.J. Simulation of microwave brightness temperatures of an evolving hailstorm at SSM/I frequencies. *Bull. Am. Meteorol. Soc.* **1990**, *71*, 2–13. [CrossRef]

27. Wilheit, T.; Adler, R.; Avery, S.; Barrett, E.; Bauer, P.; Berg, W.; Chang, A.; Ferriday, J.; Grody, N.; Goodman, S.; et al. Algorithms for the retrieval of rainfall from passive microwave measurements. *Remote Sens. Rev.* **1994**, *11*, 163–194. [CrossRef]

28. Weng, F.; Grody, N.C. Retrieval of Ice cloud parameters using a microwave imaging radiometer. *J. Atmos. Sci.* **2000**, *57*, 1069–1081. [CrossRef]

29. Bennartz, R.; Petty, G.W. The sensitivity of microwave remote sensing observations of precipitation to ice particle size distributions. *J. Appl. Meteorol.* **2001**, *40*, 345–364. [CrossRef]

30. Bauer, P.; Moreau, E.; Di Michele, S. Hydrometeor retrieval accuracy using microwave window and sounding channel observations. *J. Appl. Meteorol.* **2005**, *44*, 1016–1032. [CrossRef]

31. Kidd, C.; Kniveton, D.R.; Todd, M.C.; Bellerby, T.J. Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeor.* **2003**, *4*, 1088–1104. [CrossRef]

32. Marzano, F.S.; Palmacci, M.; Cimini, D.; Giuliani, G.; Turk, F.J. Multivariate statistical integration of Satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1018–1032. [CrossRef]

33. Kidd, C.; Huffman, G. Global precipitation measurement: Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]

34. Tapiador, F.J.; Turk, F.J.; Petersen, W.; Hou, A.Y.; García-Ortega, E.; Machado, L.A.T.; Angelis, C.F.; Salio, P.; Kidd, C.; Huffman, G.J.; et al. Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.* **2012**, *104–105*, 70–97. [CrossRef]

35. 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. Hydrometeor.* **2007**, *8*, 38–55. [CrossRef]

36. Gebremichael, M.; Hossain, F. (Eds.) *Satellite Rainfall Applications for Surface Hydrology*; Springer: Dordrecht, The Netherlands; New York, NY, USA, 2010; ISBN 978-90-481-2914-0.

37. 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. Hydrometeor.* **2004**, *5*, 487–503. [CrossRef]

38. Kubota, T.; Aonashi, K.; Ushio, T.; Shige, S.; Takayabu, Y.N.; Kachi, M.; Arai, Y.; Tashima, T.; Masaki, T.; Kawamoto, N.; et al. Global Satellite Mapping of Precipitation (GSMaP) products in the GPM Era. In *Satellite Precipitation Measurement*; Advances in Global Change Research; Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 67, pp. 355–373. ISBN 978-3-030-24567-2.

39. Hsu, K.-L.; Karbalee, N.; Braithwaite, D. Improving PERSIANN-CCS Using Passive Microwave Rainfall Estimation. In *Satellite Precipitation Measurement*; Advances in Global Change Research; Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 67, pp. 375–391. ISBN 978-3-030-24567-2.

40. Kuligowski, R.J.; Li, Y.; Hao, Y.; Zhang, Y. Improvements to the GOES-R Rainfall Rate Algorithm. *J. Hydrometeorol.* **2016**, *17*, 1693–1704. [CrossRef]

41. Hsu, K.; Gao, X.; Sorooshian, S.; Gupta, H.V. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.* **1997**, *36*, 1176–1190. [CrossRef]

42. Hsu, K.; Gupta, H.V.; Gao, X.; Sorooshian, S. Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resour. Res.* **1999**, *35*, 1605–1618. [CrossRef]

43. Sorooshian, S.; Hsu, K.-L.; Gao, X.; Gupta, H.V.; Imam, B.; Braithwaite, D. Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2035–2046. [CrossRef]

44. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.-L.; Joyce, R.J.; Kidd, C.; Nelkin, E.J.; Sorooshian, S.; Stocker, E.F.; Tan, J.; et al. Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In *Satellite Precipitation Measurement*; Advances in Global Change Research; Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 67, pp. 343–353. ISBN 978-3-030-24567-2.

45. Adler, R.F.; Negri, A.J.; Keehn, P.R.; Hakkarinen, I.M. Estimation of monthly rainfall over japan and surrounding waters from a combination of low-orbit microwave and geosynchronous IR data. *J. Appl. Meteorol.* **1993**, *32*, 335–356. [CrossRef]

46. Levizzani, V.; Porcú, F.; Marzano, F.S.; Mugnai, A.; Smith, E.A.; Prodi, F. Investigating a SSM/I microwave algorithm to calibrate Meteosat infrared instantaneous rainrate estimates. *Meteorol. Appl.* **1996**, *3*, 5–17. [CrossRef]

47. Porcú, F.; Borga, M.; Prodi, F. Rainfall estimation by combining radar and infrared satellite data for nowcasting purposes. *Meteorol. Appl.* **1999**, *6*, 289–300. [CrossRef]

48. Bellerby, T.; Todd, M.; Kniveton, D.; Kidd, C. Rainfall estimation from a combination of TRMM precipitation radar and GOES multispectral satellite imagery through the use of an artificial neural network. *J. Appl. Meteorol.* **2000**, *39*, 2115–2128. [CrossRef]

49. Todd, M.C.; Kidd, C.; Kniveton, D.; Bellerby, T.J. A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall. *J. Atmos. Ocean. Technol.* **2001**, *18*, 742–755. [CrossRef]

50. Turk, F.J.; Miller, S.D. Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1059–1069. [CrossRef]

51. Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Hirose, M.; Takayabu, Y.N.; Ushio, T.; Nakagawa, K.; et al. Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2259–2275. [CrossRef]

52. Bellerby, T.; Hsu, K.; Sorooshian, S. LMODEL: A Satellite Precipitation Methodology Using Cloud Development Modeling. Part I: Algorithm Construction and Calibration. *J. Hydrometeorol.* **2009**, *10*, 1081–1095. [CrossRef]

53. Kidd, C.; Bauer, P.; Turk, J.; Huffman, G.J.; Joyce, R.; Hsu, K.-L.; Braithwaite, D. Intercomparison of high-resolution precipitation products over Northwest Europe. *J. Hydrometeorol.* **2012**, *13*, 67–83. [CrossRef]

54. Di Paola, F.; Casella, D.; Dietrich, S.; Mugnai, A.; Ricciardelli, E.; Romano, F.; Sanò, P. Combined MW-IR Precipitation Evolving Technique (PET) of convective rain fields. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 3557–3570. [CrossRef]

55. Porcú, F.; Caracciolo, C.; Prodi, F. Cloud systems leading to flood events in Europe: An overview and classification. *Meteorol. Appl.* **2003**, *10*, 217–227. [CrossRef]

56. Miglietta, M.M.; Laviola, S.; Malvaldi, A.; Conte, D.; Levizzani, V.; Price, C. Analysis of tropical-like cyclones over the Mediterranean Sea through a combined modeling and satellite approach: TLC ANALYSIS THROUGH A COMBINED APPROACH. *Geophys. Res. Lett.* **2013**, *40*, 2400–2405. [CrossRef]

57. Manzato, A.; Davolio, S.; Miglietta, M.M.; Pucillo, A.; Setvák, M. 12 September 2012: A supercell outbreak in NE Italy? *Atmos. Res.* **2015**, *153*, 98–118. [CrossRef]

58. Panegrossi, G.; Casella, D.; Dietrich, S.; Marra, A.C.; Sano, P.; Mugnai, A.; Baldini, L.; Roberto, N.; Adirosi, E.; Cremonini, R.; et al. Use of the GPM constellation for monitoring heavy precipitation events over the Mediterranean Region. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2733–2753. [CrossRef]

59. Roberto, N.; Adirosi, E.; Baldini, L.; Casella, D.; Dietrich, S.; Gatlin, P.; Panegrossi, G.; Petracca, M.; Sanò, P.; Tokay, A. Multi-sensor analysis of convective activity in central Italy during the HyMeX SOP 1.1. *Atmos. Meas. Tech.* **2016**, *9*, 535–552. [CrossRef]

60. Silvestro, F.; Rebora, N.; Giannoni, F.; Cavallo, A.; Ferraris, L. The flash flood of the Bisagno Creek on 9th October 2014: An "unfortunate" combination of spatial and temporal scales. *J. Hydrol.* **2016**, *541*, 50–62. [CrossRef]

61. Silvestro, F.; Rebora, N.; Rossi, L.; Dolia, D.; Gabellani, S.; Pignone, F.; Trasforini, E.; Rudari, R.; De Angeli, S.; Masciulli, C. What if the 25 October 2011 event that struck Cinque Terre (Liguria) had happened in Genoa, Italy? Flooding scenarios, hazard mapping anddamage estimation. *Nat. Hazards Earth Syst. Sci.* **2016**, *16*, 1737–1753. [CrossRef]

62. Marra, A.C.; Porcù, F.; Baldini, L.; Petracca, M.; Casella, D.; Dietrich, S.; Mugnai, A.; Sanò, P.; Vulpiani, G.; Panegrossi, G. Observational analysis of an exceptionally intense hailstorm over the Mediterranean area: Role of the GPM Core Observatory. *Atmos. Res.* **2017**, *192*, 72–90. [CrossRef]

63. Puca, S.; Porcu, F.; Rinollo, A.; Vulpiani, G.; Baguis, P.; Balabanova, S.; Campione, E.; Ertürk, A.; Gabellani, S.; Iwanski, R.; et al. The validation service of the hydrological SAF geostationary and polar satellite precipitation products. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 871–889. [CrossRef]

64. Friedrich, K.; Hagen, M.; Einfalt, T. A Quality control concept for radar reflectivity, polarimetric parameters, and Doppler Velocity. *J. Atmos. Ocean. Technol.* **2006**, *23*, 865–887. [CrossRef]

65. Joss, J.; Lee, R. The Application of Radar–gauge comparisons to operational precipitation profile corrections. *J. Appl. Meteorol.* **1995**, *34*, 2612–2630. [CrossRef]

66. Germann, U.; Joss, J. Mesobeta profiles to extrapolate radar precipitation measurements above the alps to the ground level. *J. Appl. Meteorol.* **2002**, *41*, 542–557. [CrossRef]

67. Carey, L.D.; Rutledge, S.A.; Ahijevych, D.A.; Keenan, T.D. Correcting propagation effects in C-Band polarimetric radar observations of tropical convection using differential propagation phase. *J. Appl. Meteorol.* **2000**, *39*, 1405–1433. [CrossRef]

68. Testud, J.; Le Bouar, E.; Obligis, E.; Ali-Mehenni, M. The Rain Profiling Algorithm Applied to Polarimetric Weather Radar. *J. Atmos. Ocean. Technol.* **2000**, *17*, 332–356. [CrossRef]

69. Vulpiani, G.; Tabary, P.; Parent du Chatelet, J.; Marzano, F.S. Comparison of advanced radar polarimetric techniques for operational attenuation correction at C Band. *J. Atmos. Ocean. Technol.* **2008**, *25*, 1118–1135. [CrossRef]

70. Crisologo, I.; Vulpiani, G.; Abon, C.C.; David, C.P.C.; Bronstert, A.; Heistermann, M. Polarimetric rainfall retrieval from a C-Band weather radar in a tropical environment (The Philippines). *Asia-Pacific J. Atmos. Sci.* **2014**, *50*, 595–607. [CrossRef]

71. Rinollo, A.; Vulpiani, G.; Puca, S.; Pagliara, P.; Kaňák, J.; Lábó, E.; Okon, L.; Roulin, E.; Baguis, P.; Cattani, E.; et al. Definition and impact of a quality index for radar-based reference measurements in the H-SAF precipitation product validation. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2695–2705. [CrossRef]

72. Petracca, M.; D'Adderio, L.P.; Porcù, F.; Vulpiani, G.; Sebastianelli, S.; Puca, S. Validation of GPM Dual-Frequency Precipitation Radar (DPR) Rainfall Products over Italy. *J. Hydrometeorol.* **2018**, *19*, 907–925. [CrossRef]

73. Sebastianelli, S.; Russo, F.; Napolitano, F.; Baldini, L. On precipitation measurements collected by a weather radar and a rain gauge network. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 605–623. [CrossRef]

74. Tabary, P. The new french operational radar rainfall product. Part I: Methodology. *Weather Forecasting* **2007**, *22*, 393–408. [CrossRef]

75. Martins Costa do Amaral, L.; Barbieri, S.; Vila, D.; Puca, S.; Vulpiani, G.; Panegrossi, G.; Biscaro, T.; Sanò, P.; Petracca, M.; Marra, A.; et al. Assessment of Ground-reference data and validation of the H-SAF precipitation products in Brazil. *Remote Sens.* **2018**, *10*, 1743. [CrossRef]

76. Vulpiani, G.; Baldini, L.; Roberto, N. Characterization of Mediterranean hail-bearing storms using an operational polarimetric X-band radar. *Atmos. Meas. Tech.* **2015**, *8*, 4681–4698. [CrossRef]

77. Schmetz, J.; Pili, P.; Tjemkes, S.; Just, D.; Kerkmann, J.; Rota, S.; Ratier, A. Supplement to an introduction to Meteosat Second Generation (MSG): SEVIRI CALIBRATION. *Bull. Amer. Meteor. Soc.* **2002**, *83*, 992. [CrossRef]

78. Mugnai, A.; Casella, D.; Cattani, E.; Dietrich, S.; Laviola, S.; Levizzani, V.; Panegrossi, G.; Petracca, M.; Sanò, P.; Di Paola, F.; et al. Precipitation products from the hydrology SAF. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 1959–1981. [CrossRef]

79. Turk, F.J.; Rohaly, G.; Hawkins, J.; Smith, E.A.; Marzano, F.S.; Mugnai, A.; Levizzani, V. Meteorological applications of precipitation estimation from combined SSM/I, TRMM and geostationary satellite data. In *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere*; Pampaloni, P., Paloscia, S., Eds.; VSP International Science Publisher: Florence, Italy, 2000; pp. 353–363.

80. Turk, F.J.; Sohn, B.-J.; Oh, H.-J.; Ebert, E.E.; Levizzani, V.; Smith, E.A. Validating a rapid-update satellite precipitation analysis across telescoping space and time scales. *Meteorol. Atmos. Phys.* **2009**, *105*, 99–108. [CrossRef]

81. Pignone, F.; Rebora, N.; Silvestro, F.; Castelli, F. *GRISO (Generatore Random di Interpolazioni Spaziali da Osservazioni Incerte)-Piogge*; 2010; p. 353.

82. Feidas, H.; Porcu, F.; Puca, S.; Rinollo, A.; Lagouvardos, C.; Kotroni, V. Validation of the H-SAF precipitation product H03 over Greece using rain gauge data. *Appl. Climatol.* **2018**, *131*, 377–398. [CrossRef]

83. Hersbach, H.; de Rosnay, P.; Bell, B.; Schepers, D.; Simmons, A.; Soci, C.; Abdalla, S.; Alonso-Balmaseda, M.; Balsamo, G.; Bechtold, P.; et al. Operational Global Reanalysis: Progress, Future Directions and Synergies with NWP. 2018. Available online: https://www.ecmwf.int/node/18765 (accessed on 30 July 2020).

84. Berrisford, P.; Dee, D.P.; Poli, P.; Brugge, R.; Fielding, M.; Fuentes, M.; Kallberg, P.W.; Kobayashi, S.; Uppala, S.M.; Simmons, A. The ERA-Interim Archive Version 2.0. 2011. Available online: https://www.ecmwf.int/node/8174 (accessed on 30 July 2020).

85. 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]

86. Zhao, F.; Tang, C.; Dai, C.; Wu, X.; Wei, H. The Global Distribution of Cirrus Clouds Reflectance Based on MODIS Level-3 Data. *Atmosphere* **2020**, *11*, 219. [CrossRef]

87. Lynch, D. (Ed.) *Cirrus*; Oxford University Press: Cambridge, UK; New York, NY, USA, 2002; ISBN 978-0-19-513072-0.

88. Nurmi, P. Recommendations on the Verification of local Weather Forecasts. Available online: https://www.ecmwf.int/node/11401 (accessed on 30 July 2020).

Hydrometeor Distribution and Linear Depolarization Ratio in Thunderstorms

Zbyněk Sokol [1,*], Jana Minářová [1,2] and Ondřej Fišer [1,3]

[1]   Institute of Atmospheric Physics of the Czech Academy of Sciences, Bocni II, 141 00 Praha 4, Czech Republic; jana.minarova@ufa.cas.cz (J.M.); ondrej@ufa.cas.cz (O.F.)

[2]   Faculty of Science, Charles University, Albertov 6, 128 00 Praha 2, Czech Republic

[3]   Faculty of Electrical Engineering and Informatics, University of Pardubice, 532 10 Pardubice 2, Czech Republic

*   Correspondence: sokol@ufa.cas.cz; Tel.: +420-272-016-037

Abstract: The distribution of hydrometeors in thunderstorms is still under investigation as well as the process of electrification in thunderclouds leading to lightning discharges. One indicator of cloud electrification might be high values of the Linear Depolarization Ratio (LDR) at higher vertical levels. This study focuses on LDR values derived from vertically pointing cloud radars and the distribution of five hydrometeor species during 38 days with thunderstorms which occurred in 2018 and 2019 in Central Europe, close to our radar site. The study shows improved algorithms for de-aliasing, the derivation of vertical air velocity and the classification of hydrometeors in clouds using radar data. The comparison of vertical profiles with observed lightning discharges in the vicinity of the radar site (≤1 km) suggested that cloud radar data can indirectly identify "lightning" areas by high LDR values observed at higher gates due to the alignment of ice crystals, likely because of an intensified electric field in thunderclouds. Simultaneously, the results indicated that at higher gates, there is a mixture of several hydrometeor species, which suggests a well-known electrification process by collisions of hydrometeors.

Keywords: cloud radar; thunderstorm; LDR; hydrometeor; hydrometeor classification; lightning; discharge

1. Introduction

Investigation of atmospheric electricity and lightning has started several hundred years ago and intense attention to processes of cloud electrification has been examined over the last several decades. However, our knowledge is still not complete because of our limited abilities to measure and observe processes, which occur in the atmosphere. It is supposed that the existence and development of electric field in the atmosphere is related to cosmic rays [1,2] and synergy of hydrometeors in clouds [3]. Currently, it is widely accepted that the main process leading to cloud electrification and lightning discharges is the process of riming electrification, often called the non-inductive charging [3–6]. In the non-inductive charging, it is assumed that the charging occurs mainly due to collisions of hydrometeors; between ice and graupel hydrometeors in particular [3,5].

Except for mathematical models, research data on electrification processes are available from laboratory experiments [7,8] and field campaigns carried out in thunderclouds (e.g., balloon experiments, aircraft data) [9–13] or they can be derived from satellite or radar observations [14–19]. Cloud radars represent an important data source for estimation of distribution of hydrometeors and for derivation of vertical air velocity. Thereby in thunderclouds, the cloud radars may provide necessary information for the charging mechanism by collisions of hydrometeors. In addition, measurements from polarimetric cloud radars can be used to indicate the electric field in cloud. Vonnegut [20] was the first, who described that ice crystals align within the electrostatic field in thunderstorms. Since that time, there were other

studies suggesting it as well [21–24]. This alignment of highly asymmetric ice particles is assumed to be indicated by clearly higher values of depolarization [25–28].

In this paper, we study differences in distribution of hydrometeors and in values of Linear Depolarization Ratio (LDR) in thunderclouds in dependence on whether a lightning discharge was recorded in the vicinity of the radar site or not. Based on data from a vertically-oriented polarimetric cloud radar, we estimate 5 hydrometeor species and we compare the identified hydrometeor species together with LDR values with lightning observations recorded by EUCLID (European Cooperation for Lightning Detection) network.

The paper is organized as follows. After this introductory section, Section 2 provides the reader with description of the cloud radar and of algorithms, which we apply to derive vertical air velocity (AV) and to classify hydrometeor species (Hclass). This section also provides an overview of analyzed thunderstorms and describes methods of comparison between obtained or derived data from the cloud radar and recorded lightning discharges near the radar site. Section 3 displays results; it details a thunderstorm that occurred on 10 June 2019 from diverse perspectives and then it shows common characteristics and average vertical profiles of LDR of (all) analyzed thunderstorms. Section 4 discusses the obtained results, while Section 5 draws conclusions of this study.

2. Materials and Methods

2.1. Vertically-Oriented Cloud Radar

A vertically-oriented Ka-band cloud radar MIRA 35c manufactured by METEK Gmbh (http://metek.de/) was mounted at the top of Milešovka hill (837 m a. s. l.) at a meteorological observatory (Figure 1) in Central Europe (northwestern Czechia; 50°33′18″ N and 13°55′54″ E) in March 2018. The polarimetric cloud radar (i.e., cloud profiler) works at a frequency of 35 GHz. After calibration of the instrument, the radar began operating in June 2018. Table 1 provides the reader with the technical specifications of the radar.
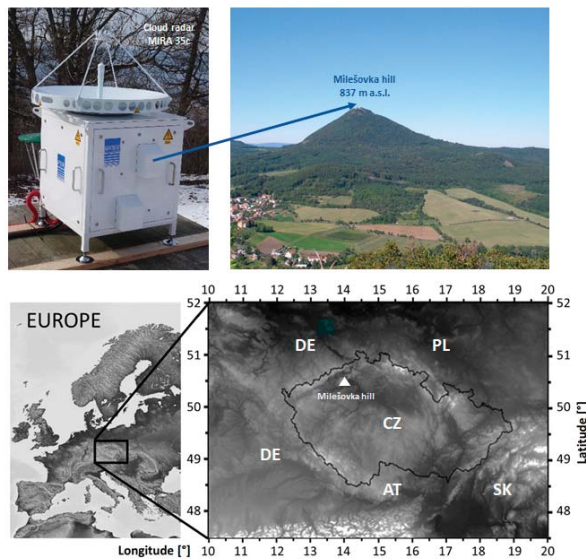
Figure 1. Placement of the cloud radar MIRA35c at the Milešovka observatory situated at the top of Milešovka hill (837 m a.s.l.) and its geographical location in Czechia in Central Europe.

Table 1. Technical parameters of the cloud radar MIRA35c situated at the Milešovka hill.

Technical Parameter	Cloud Radar MIRA 35c
Radar system	Doppler polarimetric
Radar band	Ka-band
Radar core	Magnetron type
Antenna type	Cassegrain
Transmitter frequency	35.12 GHz +/−0.1 GHz
Peak power	2.5 kW
Antenna diameter	1 m
Antenna gain	48.5 dB
Antenna beam width	0.6°
Pulse repetition frequency	2.5–10 kHz
Pulse width	min. 0.1 µs
	max. 0.4 µs
Detection unambiguous velocity range (± $V_{Nyquist}$)	±10.65 m/s
Original measurements	Doppler spectra

The cloud radar is equipped with an Interactive Data Language (IDL) software enabling the basic processing and visualization of Doppler spectra (http://metek.de/product/mira-35c/). Doppler spectra are obtained after averaging 40 consecutive values and after estimation of noise floor. Values below the estimated noise floor are considered as NaN values (they have no signal). The IDL software calculates three moments of averaged Doppler spectra such as radar reflectivity (Z), Doppler vertical velocity (DV) and spectrum width (σ). Among derived quantities, one also obtains values of Signal-to-Noise Ratio (SNR) and values of Linear Depolarization Ratio (LDR). Approximately, the radar records are available every 2 s from 509 gates. Gates are denoted ig = 4, . . . , ig = 512 because the first three gates (ig = 1, ig = 2, and ig = 3), which are closest to the ground, are not processed. The distance between two consecutive gates is 28.8 m.

In this study, we use averaged noise-free Doppler spectra to estimate vertical air velocity (AV) and derive five hydrometeor species. The algorithm deriving AV as well as the algorithm classifying hydrometeors (Hclass) are described in Sections 2.2 and 2.3.

2.2. Calculation of AV

The calculation of AV is crucial for classification of hydrometeor species (Hclass, Section 2.3) because the classification is mainly based on their terminal velocity (TV). The cloud radar measures composed velocity (DV) of TV and air velocity (AV), such as DV = TV + AV. When calculating AV, the AV is oriented towards the radar (downward) in accordance with basic processing of measured data by the IDL software used by the radar manufacturer. At the end of our calculation of AV, however, the AV orientation is reversed and we present all outputs of AV with upward orientation.

AV calculation is based on very small particles in Doppler spectra that are assumed to be so light that their TV is very close to zero, i.e., they are carried solely by air, and thus their velocity defines the AV. This is a common approach that was detailed by Kollias et al. [29], Gossard [30] and Shupe et al. [31], and conducted by, e.g., Zheng et al. [32] or Sokol et al. [33]. In this study, we innovated the algorithm used by Sokol et al. [33] to derive AV from variances caused by turbulence, wind shear, particle size distribution, and finite radar beam width.

We found that the original algorithm used for AV calculation, which performs de-aliasing of the Doppler spectra, can lead in some cases to significant and unrealistic temporal changes of AV, which then result in erroneous hydrometeor classification. That is why our new de-aliasing algorithm uses three methods of AV calculation, compares the result of each of them (AV1, AV2 and AV3) with the result of the two others and also compares the three calculated AV with the AV calculated for the previous cloud radar recording, i.e., 2 s back in time, approximately (AVL).

Any Doppler spectrum is stored for each gate in intervals beginning with a component ia and ending with a component ib, where ia corresponds to lower speed and ib to higher speed. The intervals,

which are identified by the algorithm of the manufacturer of the cloud radar, represent continuous parts of a Doppler spectrum which are ordered from the lowest speeds. Note that for a gate, we can obtain multiple intervals. Components, which are not part of any determined interval, are considered to have zero amplitude. In this study, we consider not only the interval with the lowest magnitude of velocity corresponding to ia in the first interval (Sokol, Z. et al. [33]), but also the ia from the second interval.

We assume that measured values might be aliased. Therefore, in addition to recorded values V_{ori}, we also consider $V_{ori} \pm V_d$ values, where $V_d = V_{Nyquist} + V_{Nyquist}$ (for $V_{Nyquist}$ value see Table 1). We use parameters $q_{tol} = 3$ and $q_{max1} = 5$ in the following calculations of AV. The calculations of AV consist of steps provided below, which are performed for individual gates (ig) from the bottom (ig = 4) to the top (ig = 512) because AV is not affected by aliasing in the lowest gates as updrafts cannot be that strong ($>V_{Nyquist}$) near the ground. We calculate AV1, AV2, AV3 and AVL and assign the resulting AV to the value (i.e., among AV1, AV2, AV3 and AVL) that best corresponds to conditions described in the next paragraph. In the de-aliasing algorithm, we also use a reference value V_{ref}, which is define in steps 7, 8 and 9 in the procedure described below.

The procedure of AV calculation consists of following steps:

1. For an ig (ig = 4 at first), we define de-aliasing function (DAL) calculating velocity V_{cor} using original and reference velocities V_{ori} and V_{ref}, respectively:

$$V_{cor} = DAL(V_{ori}, V_{ref}), \tag{1}$$

   where V_{cor} corresponds to one of the values (V_{ori}, $V_{ori} + V_d$ or $V_{ori} - V_d$) that is closest to the value V_{ref}.

2. AV1 is calculated in the same way as by Sokol et al. [33], i.e., we consider only those components of Doppler spectra whose amplitude is at least 0.1% of the maximum Doppler spectra amplitude.

3. AV2 is calculated in the same way as in point 2 but without the mentioned condition concerning the maximum Doppler spectra amplitude.

4. In addition to the (first) interval corresponding to the lowest speed, we also consider the (second) interval corresponding to the second lowest speed. We determine AV3 as the most left point of the second interval. The reason is that it can happen that the first interval containing the lowest speed is very narrow and far from the second interval. A closer comparison commonly shows that the first interval is likely an unremoved noise since its values are usually inconsistent with values in a gate below and a gate above and these values are also inconsistent with the values recorded in previous measurements. In such a case, it is evident that the second interval should also be considered in AV calculation (AV3).

5. We take into account AV calculated for the same gate but in the previous recording (≈ 2 s prior to the investigated recording) and we denote the resulting value AVL.

6. If AV(ig − 1) is available, then we calculate AV(ig) using the following:
   $AV1_{cor}(ig) = DAL(AV1(ig), AV(ig − 1))$
   $AV2_{cor}(ig) = DAL(AV2(ig), AV(ig − 1))$
   $AV3_{cor}(ig) = DAL(AV3(ig), AV(ig − 1))$
   If $|AV2_{cor}(ig) − AV(ig − 1)| < q_{tol}$, then AV(ig) = $AV2_{cor}(ig)$, stop
   If $|AV1_{cor}(ig) − AV(ig − 1)|$, then AV(ig) = $AV1_{cor}(ig)$, stop
   If $|AV3_{cor}(ig) − AV(ig − 1)|$, then AV(ig) = $AV3_{cor}(ig)$, stop

7. If all the above given conditions (no. 6) are fulfilled, then AV(ig) is assigned according to the nearest value among $AV1_{cor}(ig)$, $AV2_{cor}(ig)$ and $AV3_{cor}(ig)$ to V_{ref} = AV(ig − 1) if AV(ig − 1) is available.

8. If AV(ig − 1) is not available but AVL(ig) is available, then: V_{ref} = AVL(ig) and AV(ig) = $DAL(AV2(ig), V_{ref}(ig))$.

9. If AVL(ig) is not available, which usually does not happen, then AV(ig) from more past recordings is used as V_{ref}.

10. AV(ig = 4) is calculated using no. 7 and if AV(ig = 4) > q_{max1}, then AV(ig = 4) is set to AV(ig4) − V_d, because we do not allow large positive values of AV.
11. We repeat the procedure from no. 1 to no. 9 for the next gate, i.e., one gate higher (ig = ig + 1).
12. The procedure finishes at the highest gate (ig = 512) or at the (highest) gate, where we still recognize discrete intervals of Doppler spectra in the radar data.

It should be noted that the applied procedure determining AV based on very small particles (tracers) is limited by how well very small particles are identified in the Doppler spectra. It may happen, especially in the case of heavy rain, that smallest particles with negligible terminal velocity are not detected anymore due to extinction or that in the radar volume of some gates, there are only larger droplets likely due to size sorting. This is related to spots in thunderclouds with high LWC, which make the larger droplets arrive earlier to lower gates. However, our experience suggests that these are very rare cases and their effect is marginal.

Calculation of AV precedes the classification of hydrometeor species (Hclass) because AV cannot be neglected in summer thunderstorms that are under investigation in this paper. Thunderstorms are a convective phenomenon for which (strong) updrafts and downdrafts are typical.

2.3. Clasification of Hydrometeor Species—Hclass

Hclass performed in this paper is similar to Sokol et al. [33] and stands on the idea that the TV of diverse hydrometeor species differs. It is natural that the existence of a hydrometeor species depends on ambient air temperature and its shape can be identified by LDR (e.g., shape of ice crystals differs from that of rain). In this paper, we define five hydrometeor species; cloud water (i.e., cloud droplets), rain, graupel, hail, and ice and snow particles together. We merged ice and snow particles into one group following the recommendations provided by Sokol et al. [33], where the algorithm was not able to efficiently distinguish ice particles from snow.

Terminal velocity range of the 5 hydrometeor species between minimum terminal velocity (TVmin) and maximum terminal velocity (TVmax) is provided in Table 2. Terminal velocity ranges of hydrometeor species do not overlap and stem from values provided in COSMO numerical weather prediction model. We use these values as "standard" values for hydrometeors whenever we make model simulations over Czechia (e.g., Sokol et al. [34]). The actual TV of a target for any discrete interval of Doppler spectra corresponds to the result of subtraction of AV from DV for a given peak of the Doppler spectra.

Table 2. Terminal velocity range of hydrometeor species classified by Hclass.

Hydrometeor Specie	Terminal Velocity Range
Cloud water (C)	0.0001–0.15433 m/s
Rain (R)	0.15433–6.3384 m/s
Ice & Snow (IS)	0.0290–1.3133 m/s
Graupel (G)	1.3133–7.7747 m/s
Hail (H)	7.7747–10.0253 m/s

We define the ambient air temperature (T [°C]), which influences the presence of hydrometeor species, from ERA5 reanalyses (www.ecmwf.int). Specifically, we take temperature profiles of a grid point closest to the Milešovka observatory. ERA5 reanalyses provide us with hourly data at a horizontal resolution of 0.25° (geographical latitude) × 0.25° (geographical longitude). The use of ERA5 reanalyses differs from Sokol et al. [33], which used temperature profiles based on sounding measurements from Praha/Libuš station situated 60 km southeast from the Milešovka observatory. These sounding data are not only distant from the Milešovka observatory, which is a problem especially when investigating thunderstorms, but they are also available only at 00, 06 and 12 UTC. On the other hand, ERA5 reanalyses provide us with vertical profiles at a higher (i.e., hourly) temporal resolution and from a location much closer to the radar site (12 km). Thus, we consider ERA5 reanalyses more

suitable for this study. On the other hand, we are aware that the used temperature profiles are not accurate because the ERA5 data have a low resolution to describe temperature profiles in convective storms that differ from those of the surrounding air.

Similar to Sokol et al. [33], we use 0 °C as a temperature threshold in the Hclass algorithm and as in Sokol et al. [33], if T > 0 °C, then the Hclass provides cloud water, rain, graupel or hail. However, if T ≤ 0 °C, we determine the existence of supercooled water in the higher atmospheric (i.e., tropospheric) layers differently than Sokol et al. [33]. Sokol et al. [33] used a fixed threshold of −20 °C below which the supercooled water could not exist although in the case of convective storms, it can happen that the supercooled water is observed at much lower temperatures (even at −50 °C in some cases) due to strong updrafts and lack of time to freeze. Therefore, we modified the Hclass provided by Sokol et al. [33] and set that the supercooled cloud water can be found from 0 °C up to −40 °C (instead of previous −20 °C), which corresponds to generally accepted temperature range for the existence of supercooled water in mid-latitude summer thunderstorms. We define the existence of supercooled cloud water within 0 to −40 °C only if a condition of AV is met: (i) AV > 1 m/s if T is between 0 °C and −20 °C or (ii) AV(T) =1 − (T + 20)/5 if T is between −20 °C and −40 °C. The threshold for AV and relationships determining supercooled cloud water were obtained empirically based on subjective evaluation of their performance at various thunderstorms recorded by the cloud radar.

A limitation of our Hclass is that we did not have the means to objectively verify its results. Since any Hclass depends on selected terminal velocity range of individual hydrometeors and real measurements show ambiguity in the terminal velocity ranges, any change in terms of terminal velocity range of any hydrometeor species will affect the results of all Hclasses. This is the key source of uncertainty of Hclasses in general. Moreover, while fixing the values of given parameters (e.g., terminal velocity range), it is almost impossible to avoid subjectivity. Thus, we tested several values of parameters, compared obtained results and fixed the parameters to values that provided results closest to reality based on our experience and/or literature.

2.4. Analysed Data: Thunderstorms of 2018 and 2019

In this study, we used data of the cloud radar situated at the Milešovka observatory (Figure 1) such as Z, LDR, derived AV (Section 2.2) and classified hydrometeor species (Section 2.3) during days with lightning registered up to 20 km from the Milešovka observatory by the EUCLID network. The data covered a period from June 2018 to September 2019. The dataset consisted of 38 days of thunderstorms (Table 3). Continuous data records lasted at least 2 h and more than one thunderstorm could have occurred during a single (analyzed) day.

For the dataset of 38 days of thunderstorms (Table 3), we dispose of ground-based observations of lightning discharges by EUCLID network [35]. We obtained the data from Blitz Informationsdienst von Siemens (BLIDS) [36], which provides it to EUCLID network, for the whole territory of Czechia and its neighborhood. BLIDS uses the time of arrival (TOA) principle to locate lightning discharges. TOA principle is based on assumption that the electromagnetic field induced by a lightning discharge propagates at the speed of light from its origin in all directions. Individual receivers record TOA and the difference in TOA among receivers defines the location of the discharge.

The lightning data include information on geographical coordinates of the discharge (in WGS84), time of the discharge [ms], peak current [kA], polarity of the discharge, type of the discharge (Cloud-to-Ground CG and Cloud-to-Cloud CC), and on quality of data (binary information). The spatial accuracy of the lightning dataset was 0.6 km (median) at a confidence level of 95%, while the lightning detection efficiency was about 100% [37]. All the data used in this study were of good quality according to the binary information in the lightning dataset.

In addition, we obtained weather radar reflectivity factor data at various Constant Altitude Plan Position Indicator (CAPPI) levels for the 38 days of thunderstorms (Table 3) from the Czech radar network (CZRAD; Sokol et al. [38]) operated by the Czech Hydrometeorological Institute. CZRAD network consists of two C-band weather radars recording measurements every 5 min at a horizontal

resolution of 1 km. The closest radar is located 100 km southward from the Milešovka observatory. Main product of the two C-band weather radars is the radar-derived rain rate (R [mm/h]), which we calculate using the Z–R relationship, such as $Z=200R^{1.6}$, where Z is the radar reflectivity factor [mm^6/m^3] [39]. Furthermore, we had synoptic data from the Milešovka observatory at our disposal, and in this study, we used rain gauge measurements with a temporal resolution of 1 min.

Table 3. Date of 38 days of thunderstorms with lightning discharges recorded up to 20 km from the Milešovka observatory in 2018 (left panel) and 2019 (right panel).

Thunderstorms in 2018	Thunderstorms in 2019
2018-06-01	2019-05-20
2018-06-10	2019-05-25
2018-06-11	2019-06-06
2018-06-27	2019-06-10s
2018-06-28	2019-06-12
2018-07-05	2019-06-20
2018-07-21	2019-07-21
2018-07-28	2019-07-29
2018-08-02	2019-07-31
2018-08-03	2019-08-02
2018-08-04	2019-08-03
2018-08-08	2019-08-04
2018-08-13	2019-08-7
2018-08-17	2019-08-11
2018-08-24	2019-08-12
2018-09-21	2019-08-27
	2019-08-29
	2019-09-01

2.5. Methods of Comparison between Cloud Radar Data and Lightning Data

The comparison of cloud radar data with lightning data is not straightforward due to differences in their location. The cloud radar is a profiler, i.e., it registers a vertical profile from a particular location on the ground, whereas lightning data can be registered anywhere. Thus, in this study, we reduced the lightning dataset to lightning discharges that were located within a circular area around the Milešovka observatory, where the Milešovka observatory was situated in its centre. We considered several circular areas (with a radius of 1, 2, 3, 5, 7, 10, and 20 km) around the Milešovka observatory.

We decided to divide storms above the radar into storms with lightning (denoted "near lightning", NL) and those without lightning (denoted "far lightning", FL). After several testing, we selected 1 km as the radius defining NL because we believe that this is the distance, which describes the condition of a cloud above the Milešovka observatory, when the development of lightning (i.e., electrification) is taking place or will soon take place over the Milešovka observatory (or its nearest vicinity). Contrary, the FL (lightning registered at a distance of 10-20 km from the observatory) represent the condition of a cloud above the Milešovka observatory, when no discharges were recorded in its direct vicinity. In Section 3, we depict results for NL as compared to FL.

Another issue in comparing cloud radar data with lightning data is the difference in their temporal resolution. Cloud radar data are recorded every 2 s approximately, while lightning data are registered at a temporal resolution in the order of ms. However, this is not of high importance since based on our experience the temporal variability of radar measurements is not very high, even in thunderstorms.

Being aware of the difference in temporal resolutions of radar and lightning data, we still decided to consider cloud radar data registered just before and just after the time of any lightning discharge; i.e., we considered two consecutive cloud radar recordings (distant by 2 s) that we coupled with a lightning discharge. This way, we compared cloud radar data (vertical profiles of Hclass, AV and LDR) with lightning data in circular areas around the Milešovka observatory (NL vs. FL) during the 38 days

of thunderstorms (Table 3). Note that there were 990 lightning discharges observed up to 1 km from the Milešovka observatory during the 38 days of thunderstorms and 171,754 FL discharges.

3. Results

This section is divided into two parts. The first part describes a particular thunderstorm that occurred on 10 June 2019. We selected this thunderstorm because on that date, the observer recorded a thunder less than 1 s after he saw the lightning flash at the Milešovka observatory. According to geographical coordinates in EUCLID data, the lightning discharge occurred at a distance of only 65 m from the observatory. This lightning discharge was the second closest in our dataset; on 1 June 2018, there was a lightning detected directly at the observatory, which hit the observatory according to the book of records written by observers. Nevertheless, we do not detail this thunderstorm (on 1 June 2018) in this study since Sokol et al. [33] already studied it in detail.

Contrary to the first part of this section, which is dedicated to one particular thunderstorm, the second part of this section presents common characteristics, including LDR, that were typical throughout all thunderstorms that occurred in the 38 days in the dataset (Table 3). It presents results from the comparison of cloud radar data with lightning data in dependence on the distance of lightning discharge to the Milešovka observatory (i.e., NL vs. FL).

3.1. Thunderstorm on 10 June 2019

Figure 2 shows the temporal evolution of rain rates (as derived from C-band weather radar data, Figure 2a) and lightning discharges (Figure 2b) during one hour of the thunderstorm on 10 June 2019, when most lightning activity was observed close to the Milešovka hill. The temporal evolution is displayed with a time step of 10 min. Based on rain rates (Figure 2a), it is clearly visible that the thunderstorm was severe, at least within the central European context.

Figure 2 also shows that the lightning discharges occurred (horizontally) not only within the precipitation areas but also outside of the precipitation cores; i.e., they may have originated in non-precipitating parts of the thundercloud as well. As the system moved in time towards east-northeast, CC flashes had a tendency to precede CG flashes, while lightning flashes (CC + CG together) tended to occur not only during the period of intense rainfall, but also prior to intense rainfall. This has been mentioned in other works as well, e.g., [40].

Figure 3 depicts precipitation totals with a time step of 1 min, as registered by a rain gauge with a resolution of 0.1 mm at the Milešovka observatory between 21 and 22 h UTC on 10 June 2019. Note that rain rates (i.e., precipitation intensities, Figure 2a) cannot be directly compared with recorded precipitation totals (Figure 3) as they do not represent the same information. Figure 3 shows that the highest 1-min precipitation total occurred between 21:39 and 21:40, while the closest lightning was recorded at 21:37 and 50.341 s. This confirms that lightning flashes may occur prior to heavy rain. Here, we note that measured precipitation totals by the rain gauge might be underestimated. The reason is that at the top of Milešovka hill, strong winds frequently appear during storms, which may result in an underestimation of rainfall totals due to the blowing away of the precipitation from the surface of the rain gauge.
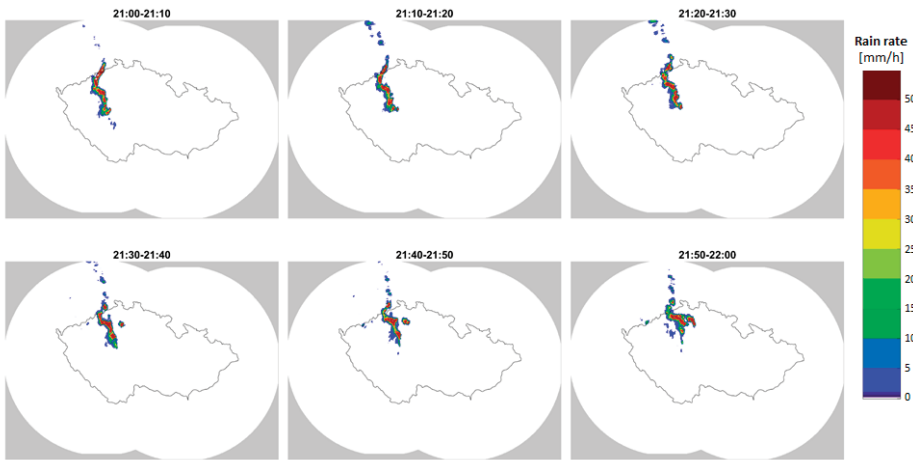
(**a**)
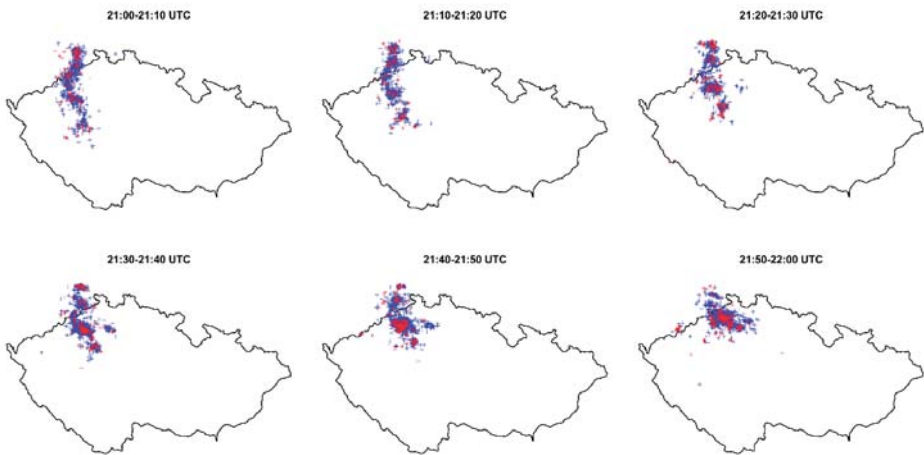
(**b**)

Figure 2. Temporal evolution with a 10-min time step of (**a**) rain rates derived from C-band weather radar data (radar circle domains having a radius of 250 km) and (**b**) lightning discharges between 21 and 22 h UTC as registered by EUCLID network on 10 June 2019. Note that in (**b**), blue symbols represent Cloud-to-Cloud (CC) flashes while red symbols display Cloud-to-Ground (CG) flashes. The Milešovka position is shown in Figure 1.

Concerning data of the cloud radar, Figure 4 displays radar reflectivity together with NL (i.e., lightning discharges that occurred not farther than 1 km from the Milešovka observatory). It is obvious from the figure that NL were related to high reflectivity values, although high reflectivity values were also typical for the melting layer and below (i.e., below 2.5 km approximately).
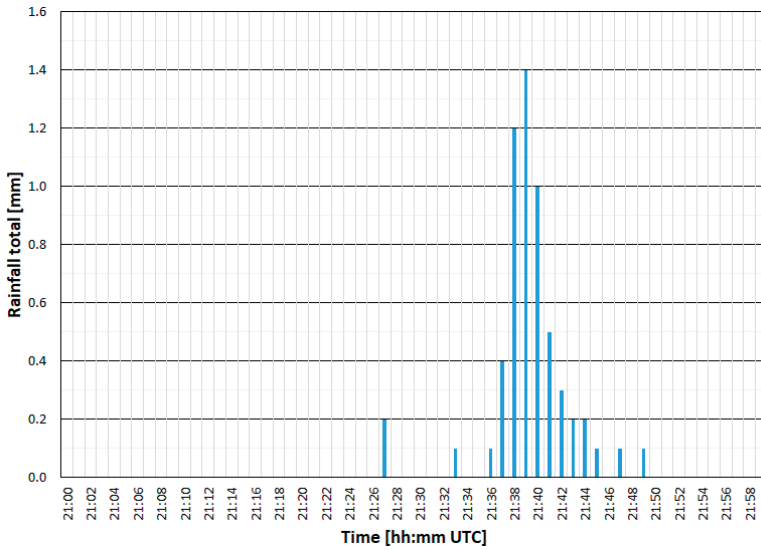
Figure 3. Rainfall totals with a 1-min time step recorded by a rain gauge at the Milešovka observatory from 21 to 22 h UTC on 10 June, 2019.
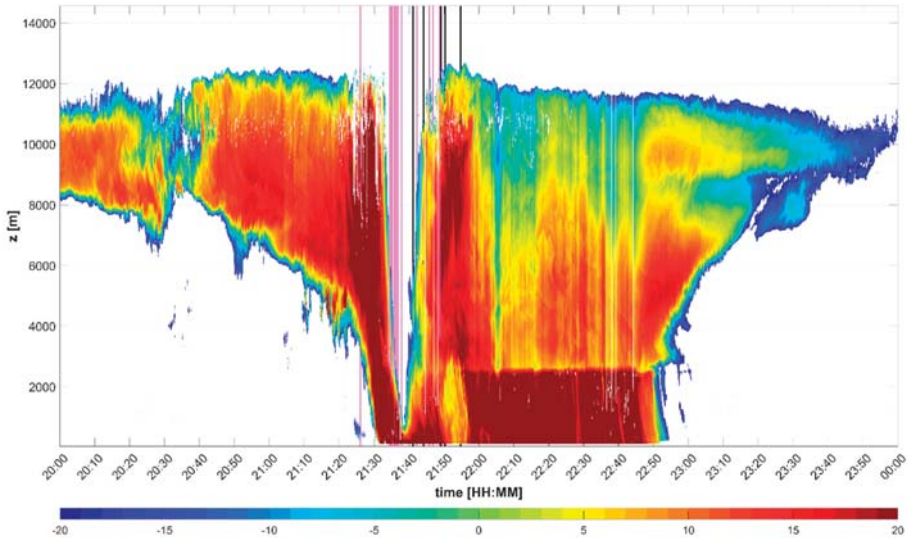
Figure 4. Temporal evolution of radar reflectivity factor (Z [dBZ]) of the cloud radar at the Milešovka observatory during the thunderstorm on 10 June 2019 from 20 to 24 h UTC. Vertical magenta lines depict CC near-lightning (NL) discharges, while vertical black lines display CG NL discharges. Note that y-axis (z [m]) displays the height in meters above the cloud radar situated at an elevation of 837 m a.s.l.

It is worthy of note that between 21:30 and 21:40 approximately, there was a sudden decrease in the vertical span of the thundercloud, according to the cloud radar data (Figure 4). This is the time when most of rainfall was registered at the Milešovka observatory (Figure 3). This is probably caused

by the fact that the received signal in the lowest gates was too strong due to heavy rain that the radar was unable to capture signal from higher gates.

Furthermore, it is interesting to check the temporal evolution of LDR values during the thunderstorm (Figure 5) that were not corrected using the integrated cross-polarization ratios [41]. In Figure 5, high values of LDR clearly show that the melting layer was around 2500 m above ground in the thunderstorm. Another zone of high LDR values is visible from 21:30 to 21:50 at higher altitudes, which is the time interval of intense rainfall (Figure 3) and lightning activity near the radar site (Figure 4). Contrary to very high LDR values in the melting layer, which are commonly associated with melting snow flakes, very high LDR values at higher altitudes, such as 4-7 km, can correspond to non-spherical shape of graupel and/or hail or to aligned ice crystals due to a strong electric field if the crystals are not aligned along with the co-channel, instead they are oriented at angle close to 45° with both the co- and cross-channels [25,42,43]. It is worthy of note that the elevation around 4-7 km, where we observed increased LDR, is also considered as the elevation where the main negatively charged area appears [40]. We discuss this finding further below.
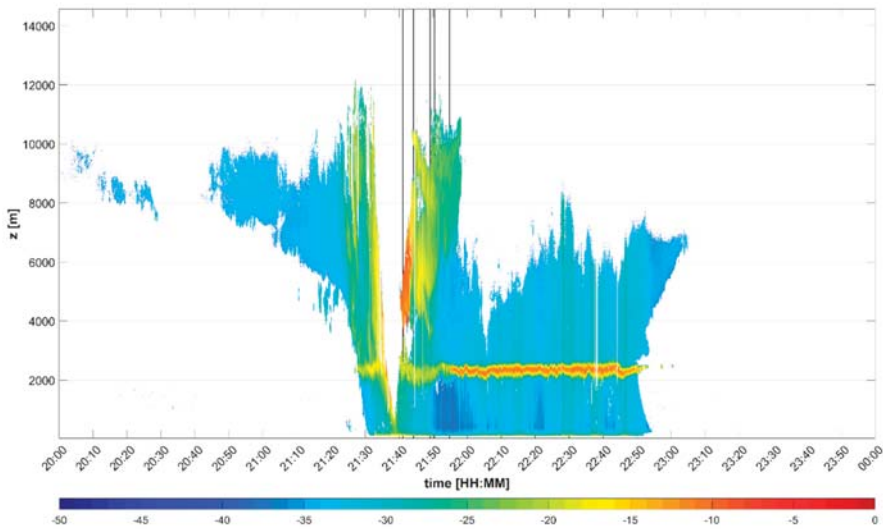
Figure 5. Temporal evolution of linear depolarization ratio (LDR) [dB] during the thunderstorm on 10 June 2019 from 20 to 24 h UTC at the Milešovka observatory. Vertical (black) lines depict CG NL discharges and y-axis (z [m]) displays the height in meters above the cloud radar situated at an elevation of 837 m a.s.l.

It should be noted that the LDR data are not available at all gates where we obtained radar reflectivity factor data (for example, after 21:50). This is the consequence of the attenuation of the signal received in the plane perpendicular to the transmission plane.

Figure 6 shows the evolution of hydrometeor distribution (resulting from Hclass, Section 2.3) on 10 June 2019 from 21:00 to 21:59 UTC. Clearly, the lack of data in the vertical profiles between approximately 21:30 and 21:40 makes the interpretation of the obtained results difficult, especially because many NL discharges occurred at that time. Nevertheless, the majority of the ten closest discharges (Figure 6) occurred after 21:40, when we had data again, covering almost all the troposphere.
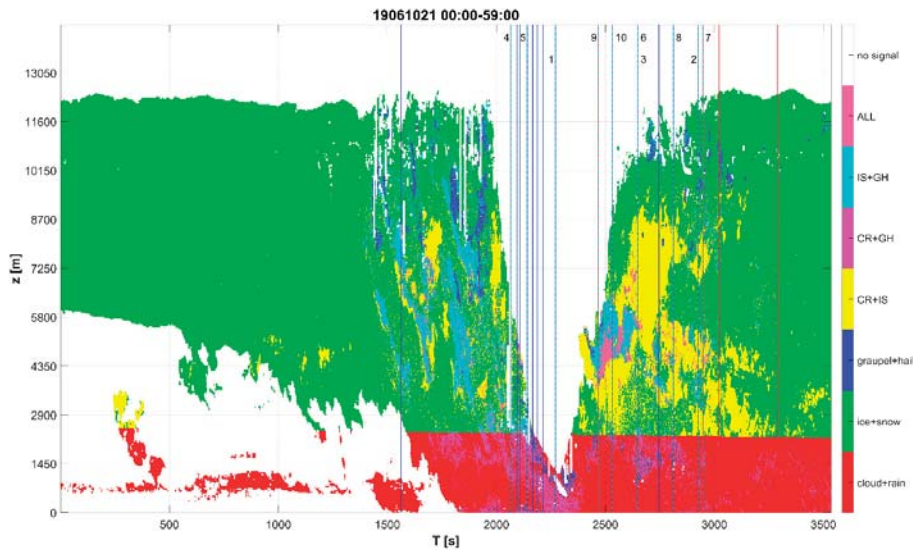
Figure 6. Hclass during the thunderstorm on 10 June 2019 from 21:00 to 21:59 UTC at the Milešovka observatory. Vertical lines depict NL discharges; blue lines depict CC NL discharges, red lines the CG NL discharges. Ten lightning discharges closest to the observatory are numbered in ascending order starting from the closest discharge denoted no. 1. Dashed lines represent cases, when more CG/CC discharges occurred around the same time (order of ms). Note that capital letters in the legend indicate first letter of classified hydrometeors (Table 2) and y-axis (z [m]) displays the height in meters above the cloud radar situated at an elevation of 837 m a.s.l.

The results of Hclass indicate that the highest LDR values at the elevation from 4 to 7 km (Figure 5) correspond to a mixture of several hydrometeor species with a predominance of ice and snow particles and graupel. These are the species which play major roles in the process of cloud electrification by collisions of hydrometeors according to currently accepted theories [44]. The mixture of many hydrometeor species is also evident during very close lightning activity (between 2400 s and almost 3000 s in Figure 6).

The electrification process by collisions of hydrometeors at an elevation of 4–7 km is also supported by Figure 7, which presents values of the Doppler spectrum width (σ). High values of σ, i.e., large variability of vertical velocities, just after 21:40 confirm the coexistence of various hydrometeors and support the existence of collisions of hydrometeors (light species collide with heavier species having larger terminal velocity). The obtained results of high LDR and sigma values together with the presence of diverse hydrometeor species may bring us to the conclusion that around 21:40, collisions of hydrometeors caused a strong electrification of the thundercloud near the radar site.

3.2. Common Characteristics of Analyzed Thunderstorms

This subsection focuses on results related tall thunderstorms in the dataset (Table 3). It shows their common (different) features and compares them with recent knowledge on lightning processes. Our intention was to compare NL with FL, when clouds were present above the observatory. Therefore, in the statistical evaluation, we used data from only those gates, where Hclass identified at least one hydrometeor species.
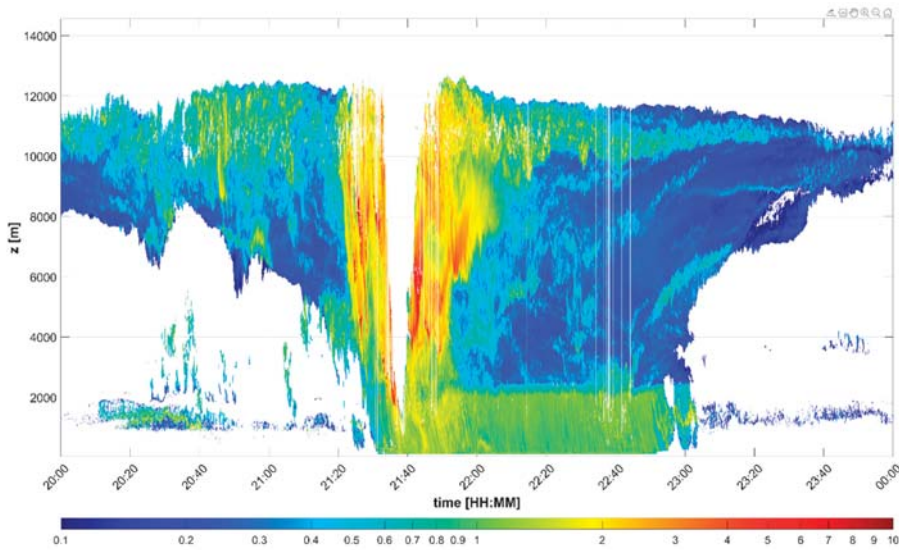
Figure 7. Doppler spectrum width (σ [m/s]) during the thunderstorm on 10 June 2019 from 20 to 24 h UTC at the Milešovka observatory. Note that the color bar is in logarithmic scale and that y-axis (z [m]) displays the height in meters above the cloud radar situated at an elevation of 837 m a.s.l.

Figure 8 summarizes the results throughout the analyzed thunderstorms at the Milešovka observatory. It depicts radar-derived quantities for NL discharges compared to that for FL discharges. It clearly shows that on average, hail, rain and graupel occurred in lower gates more frequently during NL as compared to FL. For FL, rain and graupel were almost not detected at all. For NL, hail concentration was higher at an elevation of 2000 to 2500 m above ground. This is the level which roughly corresponds to the melting layer (Figure 4). Rain concentration was higher at lower elevations, at 1800 m approximately. Thereby, it can be suggested that the closer the lightning, the higher the concentration of rain and hail. This agrees with our previous results based on 10 thunderstorms [45].

In addition, Figure 8 displays the results of AV for NL vs. FL. It shows that in the case of NL, the downward motion of the air substantially prevails at lower altitudes; from the ground to 1000 m and from 2000 to 3000 m. The layer between 1000 and 2000 m above ground is characterized by fluctuations in AV, which can be related to an interchange of up- and downdrafts. Updrafts mostly dominate the elevation from 3000 m upwards. Slow updrafts are typical for very upper vertical levels (above 9500 m).

Overall, AV seems to be quite low, which is caused by averaging. The variability in AV among gates seems high for NL. This is caused by much lower number of NL discharges (990) as compared to FL discharges (171,754), as shown in Figure 9. Concerning FL, AV does not fluctuate much on average between neighboring gates, which is due to large number of processed data. Figure 8 also shows that for FL, downward motion prevails from the ground up to 3000 m, while upward motion dominates the layers above 3000 m on average.
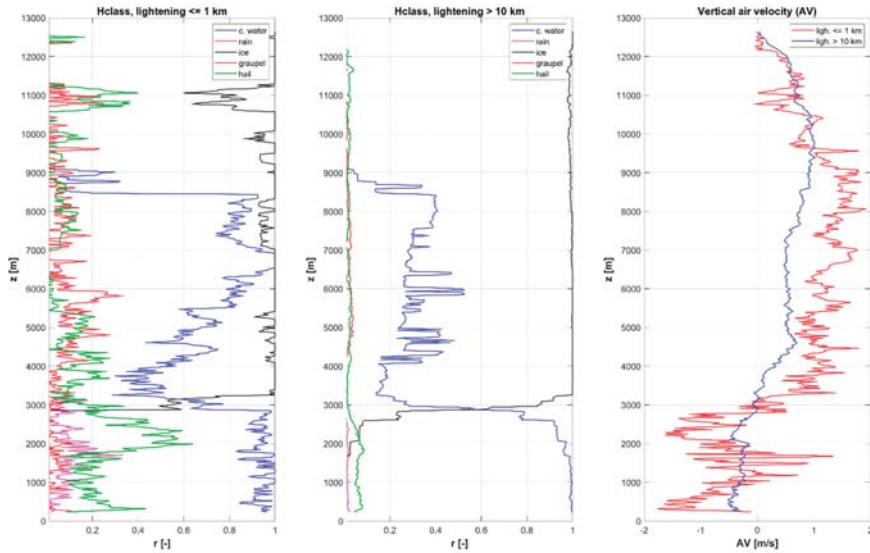
Figure 8. Vertical profile of the percentage ratio of the occurrence of hydrometeors (r [−]) during analyzed thunderstorms for: NL discharges (left panel) and far-lightning (FL) discharges (middle panel). Right panel depicts the vertical profile of air velocity (AV) oriented upwards from the cloud radar during the thunderstorms: red curve displays averages for NL discharges, while blue curve the averages for FL discharges. Note that y-axis (z [m]) represents the height in meters above the cloud radar situated at an elevation of 837 m a.s.l.
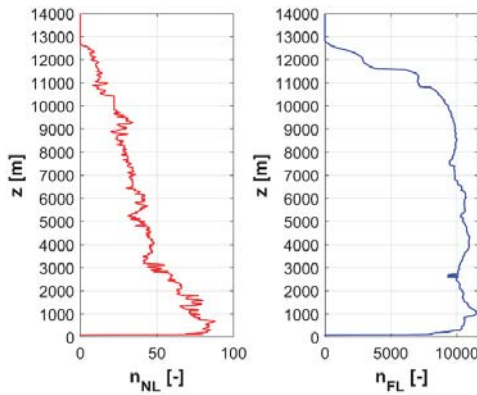
Figure 9. Number of examined cases of NL (left) and FL (right) with available LDR at gates. Y-axis represents the height [m] above the cloud radar situated at an elevation of 837 m a.s.l.

Taking into account the distance of the lightning from the observatory, we do not know whether the radar measurements took place in the frontal or back side of the thunderstorms or on their lateral sides. The placement within the thunderstorm may lead to diverse directions and values of AV, which can be confirmed by high variability of AV (not depicted). In addition to the uncertainty regarding the localization of measurements with respect to the movement and development of thunderstorms, it should be emphasized that we present results and quantities that are derived indirectly (i.e., not directly measured). Therefore, the results cannot be explicitly verified. However, we can state that

the obtained results are in accordance with the general knowledge about thunderstorms. Therefore, we believe that the technique used to calculate the vertical air velocity and to classify hydrometeors give realistic results.

3.3. LDR during Analyzed Thunderstorms

Averages of LDR are depicted in Figure 10 for NL, as compared to FL. For both NL and FL, the melting layer is not pronounced in LDR averages; there is no obvious increase in LDR averages in lowest gates. This is very likely related to the fact that the height of the melting layer depends on current atmospheric conditions, which change from one thunderstorm to another and might also change during one particular thunderstorm. As a consequence, the height of the melting layer becomes smooth in averaging, making it imperceptible in the figure.
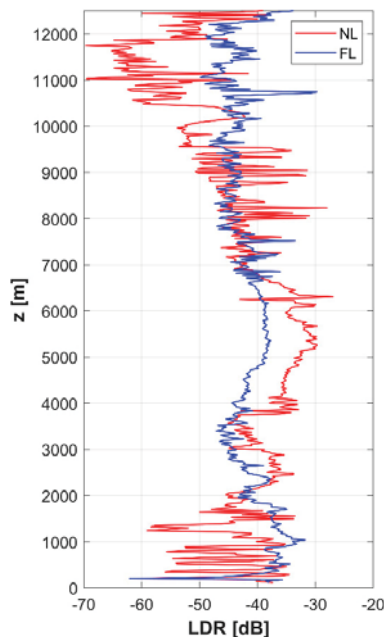
Figure 10. Vertical profile of mean LDR during thunderstorms observed at the Milešovka observatory for NL (red curve) as compared to FL (blue curve). Y-axis represents the height [m] above the cloud radar situated at an elevation of 837 m a.s.l.

The character of the curves in Figure 10 (their oscillation) is influenced by the number of averaged cases (Figure 9). This is especially true for the red curve representing NL discharges. The isolated maxima of LDR averages are probably random. However, Figure 10 clearly depicts that at an elevation of 4 to 6.5 km approximately, there are large LDR averages, which show little oscillations for NL, thus they do not correspond to random processes. These averages are much larger than the LDR averages for FL. As in Section 3.1, we attribute it to electrification by collisions and alignment of ice crystals.

To better assign the cause of increased LDR averages in the middle troposphere, Figure 11 shows 1 km layers of frequency of LDR in profiles with similar distribution of graupel and hail (i.e., rounded hydrometeors). Because concentrations of graupel and hail are similar in both NL and FL (ice or snow being present almost everywhere in these 1 km layers, Figure 8), it is obvious that there had to be another process that made the higher LDR more frequent in the case of NL as compared to FL. We suggest that the additional process could be the alignment of ice crystals observed by other

researchers, e.g., by Melnikov et al. [25]. However, we are aware that this hypothesis cannot be exactly verified.
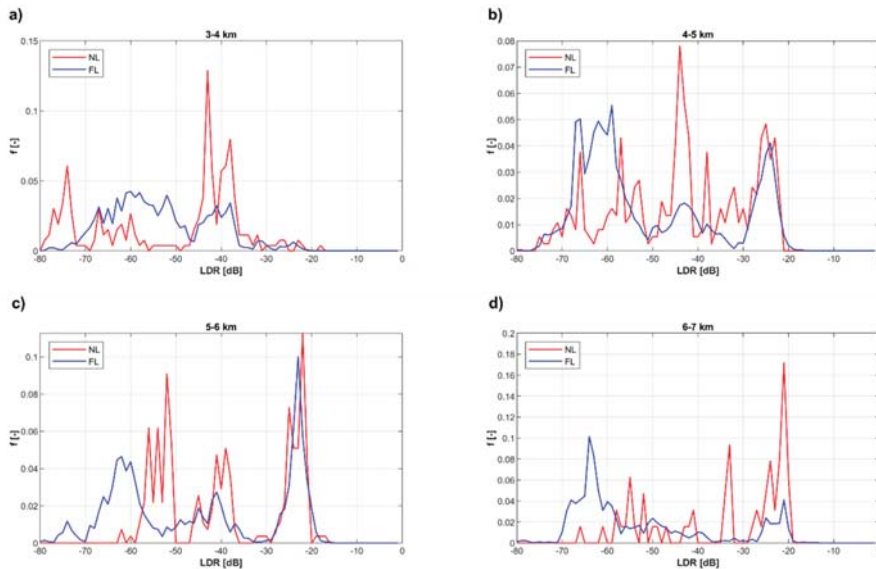
Figure 11. Frequency of LDR (f[−]) at an elevation of: (**a**) 3–4 km, (**b**) 4–5 km, (**c**) 5–6 km and (**d**) 6–7 km during thunderstorms observed at the Milešovka observatory for NL (red curve) as compared to FL (blue curve), when taking into account similar vertical profiles of graupel and hail concentrations. Y-axis represents the height [m] above the cloud radar situated at an elevation of 837 m a.s.l.

At an elevation between 8 and 9 km, significantly higher averages of LDR for NL, as compared to FL, could be rather random because of high oscillations of that for NL. On the other hand, the high oscillations of LDR averages for NL can also be related to the orientation of aligned ice crystals in an electrified field. LDR can increase if the particles align at an angle close to 45° from both the co- and cross-channels, while it can decrease if the particles align along with the co-channel (LDR reaches large negative values). Thus, the LDR of non-spherical targets, such as ice crystals, can have strongly different values (large and small) depending on the azimuth direction to the channels. Therefore, the variability of LDR may be increased in the case of NL.

The results also suggest that the clouds producing lightning in the vicinity (NL) are vertically massive and higher than clouds producing FL, at least in our dataset.

4. Discussion

The method we chose, to distinguish between NL and FL, was based on the distance of observed lightning from the radar site. The results showed that the method can be applicable in general, however, it would be better if we could also identify in which part of the storm the vertical radar measurements are taken. We plan to focus on an analysis of the possibility to identify the position of radar within storms in future. Specifically, our aim is to study whether the radar is located on the frontal or back side of a storm or on its lateral sides. Such an analysis, however, needs a wider dataset of thunderstorms registered close to the radar, which we expect to obtain in future.

At present, the amount of data during two years of radar operation that meet the condition for NL is not sufficient to allow us to divide data into further subsets, which is necessary for the identification of where—in the thunderstorm—the radar is located. The division of the current dataset into further

subsets would not lead to sufficiently robust results. Moreover, determining the position of the radar within a thunderstorm is not trivial and needs thorough investigations.

The distribution of data and the method of processing (averaging) inevitably led to a smoothing of AV. This smoothing (averaging) of vertical velocity resulted in low values of AV, since the dataset comprised both positive and negative velocities, which, after averaging, became naturally low. Nevertheless, regardless of AV smoothing, the profiles of mean AV for NL as compared to that for FL qualitatively correspond to our knowledge. During the mature state of a thunderstorm, (strong) positive as well as negative vertical velocities are supposed to be observed.

The applied technique for the recognition of hydrometeors (Hclass) provides meaningful results, although, of course, we cannot perform its exact verification. However, the structure of hydrometeors seems right during the maximum activity of the thunderstorm on 10 June 2019 (around 21:40 and later). At that time, there was a noticeable occurrence of all types of hydrometeors, including hail and supercooled cloud water throughout the vertical profile. The results obtained by processing all thunderstorms were also reasonable and explainable.

As far as the supercooled cloud water is concerned, there are no given rules on how to determine it exactly. In this study, we used a simple algorithm that allows the supercooled water to occur up to $-40\,^{\circ}\text{C}$ under the condition that there are small values in the measured Doppler spectra and there is an updraft of the air motion. We are aware that such an identification of supercooled cloud water is burdened with uncertainties. However, we consider our results satisfying.

We studied how the LDR differs for stormy areas (NL) and non-stormy areas (FL) and in our opinion, we showed, in agreement with other works e.g., [25], that a strong electric field in a thundercloud can be identified by high LDR values. This opinion is based on the fact that the increased averages of LDR at altitudes of 4 to 6.5 km were reflected for NL only. Our analysis showed that it is unlikely that the increase in LDR would be solely related to the presence of hail or graupel. On the contrary, we believe that two processes occur almost simultaneously: (i) electric field formation due to collisions of graupel and ice particles and (ii) the alignment of ice particles in the electric field leading to high LDR. It should be mentioned that aligned targets may cause characteristic signatures in the differential phase between co- and cross-channel IQ signals, but the interpretation of these signatures is difficult and we would like to address them in the future.

5. Conclusions

We investigated 38 days of thunderstorms when lightning discharges were detected in the vicinity of the Milešovka observatory; the site of a vertically-oriented Ka-band cloud radar. We analyzed vertical profiles of diverse cloud radar-derived quantities to find differences between characteristics when a lightning discharge was recorded close to the radar site ($\leq$1 km, NL) and characteristics when a lightning discharge was observed from 10-20 km from the radar site, i.e., there was a non-thunderstorm cloud above the radar (FL).

We concentrated on hydrometeor distribution, values of LDR and vertical air velocities (AV). We are aware that in most cases, we worked with data that were not directly measured; instead, we derived them (hydrometeor species, AV). The way we processed the data and derived the variables may have affected our results. Nevertheless, we believe that the procedure we chose can contribute to the current knowledge and/or the confirmation of the current knowledge on the occurrence of lightning in the atmosphere.

We showed that our technique of classifying hydrometeors provides outputs that reasonably describe thundercloud structures. Since the classification essentially depends on derived AV, our results indirectly prove that the way we derive AV provides acceptable results, although we cannot confirm it exactly.

The analysis of data characteristics of NL as compared to FL showed that:

- In the case of NL, the vertical profiles contain vertically-oriented areas with clearly high LDR, likely caused by an intensified electric field, which makes the ice particles align. The areas

with increased LDR are visible at an elevation from 4 to 6.5 km above the radar, approximately. This finding confirms results published in other studies. Unlike other studies, which usually analyze one single event, we processed 38 days of thunderstorms.

- The vertically-oriented areas with increased LDR comprise various hydrometeors, namely the ice and snow particles, graupel, hail, and (supercooled) cloud water. These are the areas which meet the condition for the development of electrification by the collision of hydrometeors. In our opinion, electric field formation due to the collisions of graupel and ice particles is followed by the alignment of ice particles in the electric field and both the processes contribute to increases in LDR.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Usoskin, I.G.; Kovaltsov, G.A. Cosmic ray induced ionization in the atmosphere: Full modeling and practical applications. *J. Geophys. Res.* **2006**, *111*, D21206. [CrossRef]
2. Sato, T. Analytical Model for Estimating Terrestrial Cosmic Ray Fluxes Nearly Anytime and Anywhere in the World: Extension of PARMA/EXPACS. *PLoS ONE* **2015**, *10*, e0144679. [CrossRef] [PubMed]
3. Saunders, C.P.R.; Bax-norman, H.; Emersic, C.; Avila, E.E.; Castellano, N.E. Laboratory studies of the effect of cloud conditions on graupel/crystal charge transfer in thunderstorm electrification. *Q. J. R. Meteorol. Soc.* **2006**, *132*, 2653–2673. [CrossRef]
4. Takahashi, T. Riming Electrification as a Charge Generation Mechanism in Thunderstorms. *J. Atmos. Sci.* **1978**, *35*, 1536–1548. [CrossRef]
5. Takahashi, T.; Miyawaki, K. Reexamination of Riming Electrification in a Wind Tunnel. *J. Atmos. Sci.* **2002**, *59*, 1018–1025. [CrossRef]
6. Takahashi, T.; Sugimoto, S.; Kawano, T.; Suzuki, K. Riming Electrification in Hokuriku Winter Clouds and Comparison with Laboratory Observations. *J. Atmos. Sci.* **2016**, *74*, 431–447. [CrossRef]
7. Saunders, C.P.R.; Peck, S.L. Laboratory studies of the influence of the rime accretion rate on charge transfer during crystal/graupel collisions. *J. Geophys. Res. Atmos.* **1998**, *103*, 13949–13956. [CrossRef]
8. Saunders, C. Charge Separation Mechanisms in Clouds. *Space Sci. Rev.* **2008**, *137*, 335–353. [CrossRef]
9. Stolzenburg, M.; Rust, W.D.; Marshall, T.C. Electrical structure in thunderstorm convective regions: 2. Isolated storms. *J. Geophys. Res.* **1998**, *103*, 14079–14096. [CrossRef]
10. MacGorman, D.R.; Biggerstaff, M.I.; Waugh, S.; Pilkey, J.T.; Uman, M.A.; Jordan, D.M.; Ngin, T.; Gamerota, W.R.; Carrie, G.; Hyland, P. Coordinated lightning, balloon-borne electric field, and radar observations of triggered lightning flashes in North Florida: Triggered lightning and storm charge. *Geophys. Res. Lett.* **2015**, *42*, 5635–5643. [CrossRef]
11. Weinheimer, A.J.; Dye, J.E.; Breed, D.W.; Spowart, M.P.; Parrish, J.L.; Hoglin, T.L.; Marshall, T.C. Simultaneous measurements of the charge, size, and shape of hydrometeors in an electrified cloud. *J. Geophys. Res. Atmos.* **1991**, *96*, 20809–20829. [CrossRef]
12. Winn, W.P.; Schwede, G.W.; Moore, C.B. Measurements of electric fields in thunderclouds. *J. Geophys. Res.* **1974**, *79*, 1761–1767. [CrossRef]

13. Weiss, S.A.; Rust, W.D.; MacGorman, D.R.; Bruning, E.C.; Krehbiel, P.R. Evolving Complex Electrical Structures of the STEPS 25 June 2000 Multicell Storm. *Mon. Weather Rev.* **2008**, *136*, 741–756. [CrossRef]

14. Adamo, C.; Goodman, S.; Mugnai, A.; Weinman, J.A. Lightning measurements from satellites and significance for storms in the mediterranean. In *Lightning: Principles, Instruments and Applications: Review of Modern Lightning Research*; Betz, H.D., Schumann, U., Laroche, P., Eds.; Springer: Dordrecht, Netherlands, 2009; pp. 309–329. ISBN 978-1-4020-9079-0.

15. Vorpahl, J.A.; Sparrow, J.G.; Ney, E.P. Satellite Observations of Lightning. *Science* **1970**, *169*, 860–862. [CrossRef] [PubMed]

16. Saha, K.; Damase, N.P.; Banik, T.; Paul, B.; Sharma, S.; De, B.K.; Guha, A. Satellite-based observation of lightning climatology over Nepal. *J. Earth Syst. Sci.* **2019**, *128*, 221. [CrossRef]

17. Sparrow, J.G.; Ney, E.P. Lightning Observations by Satellite. *Nature* **1971**, *232*, 540–541. [CrossRef]

18. Labrador, L. The detection of lightning from space. *Weather* **2017**, *72*, 54–59. [CrossRef]

19. Makowski, J.A.; MacGorman, D.R.; Biggerstaff, M.I.; Beasley, W.H. Total Lightning Characteristics Relative to Radar and Satellite Observations of Oklahoma Mesoscale Convective Systems. *Mon. Weather Rev.* **2012**, *141*, 1593–1611. [CrossRef]

20. Vonnegut, B. Orientation of Ice Crystals in the Electric Field of a Thunderstorm. *Weather* **1965**, *20*, 310–312. [CrossRef]

21. Hendry, A.; McCormick, G.C. Radar observations of the alignment of precipitation particles by electrostatic fields in thunderstorms. *J. Geophys. Res.* **1976**, *81*, 5353–5357. [CrossRef]

22. Krehbiel, P.; Chen, T.; McCrary, S.; Rison, W.; Gray, G.; Brook, M. The use of dual channel circular-polarization radar observations for remotely sensing storm electrification. *Meteorl. Atmos. Phys.* **1996**, *59*, 65–82. [CrossRef]

23. Metcalf, J.I. Radar Observations of Changing Orientations of Hydrometeors in Thunderstorms. *J. Appl. Meteor.* **1995**, *34*, 757–772. [CrossRef]

24. Biggerstaff, M.I.; Zounes, Z.; Addison Alford, A.; Carrie, G.D.; Pilkey, J.T.; Uman, M.A.; Jordan, D.M. Flash propagation and inferred charge structure relative to radar-observed ice alignment signatures in a small Florida mesoscale convective system. *Geophys. Res. Lett.* **2017**, *44*, 8027–8036. [CrossRef]

25. Melnikov, V.; Zrnić, D.S.; Weber, M.E.; Fierro, A.O.; MacGorman, D.R. Electrified Cloud Areas Observed in the SHV and LDR Radar Modes. *J. Atmos. Ocean. Technol.* **2018**, *36*, 151–159. [CrossRef]

26. Caylor, I.J.; Chandrasekar, V. Time-varying ice crystal orientation in thunderstorms observed with multiparameter radar. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 847–858. [CrossRef]

27. Ryzhkov, A.V.; Zrnić, D.S. Depolarization in Ice Crystals and Its Effect on Radar Polarimetric Measurements. *J. Atmos. Ocean. Technol.* **2007**, *24*, 1256–1267. [CrossRef]

28. Hubbert, J.C.; Ellis, S.M.; Chang, W.-Y.; Rutledge, S.; Dixon, M. Modeling and Interpretation of S-Band Ice Crystal Depolarization Signatures from Data Obtained by Simultaneously Transmitting Horizontally and Vertically Polarized Fields. *J. Appl. Meteor. Climatol.* **2014**, *53*, 1659–1677. [CrossRef]

29. Kollias, P.; Albrecht, B.A.; Lhermitte, R.; Savtchenko, A. Radar Observations of Updrafts, Downdrafts, and Turbulence in Fair-Weather Cumuli. *J. Atmos. Sci.* **2001**, *58*, 1750–1766. [CrossRef]

30. Gossard, E.E. Measurement of Cloud Droplet Size Spectra by Doppler Radar. *J. Atmos. Ocean. Technol.* **1994**, *11*, 712–726. [CrossRef]

31. Shupe, M.D.; Kollias, P.; Matrosov, S.Y.; Schneider, T.L. Deriving Mixed-Phase Cloud Properties from Doppler Radar Spectra. *J. Atmos. Ocean. Technol.* **2004**, *21*, 660–670. [CrossRef]

32. Zheng, J.; Liu, L.; Zhu, K.; Wu, J.; Wang, B. A Method for Retrieving Vertical Air Velocities in Convective Clouds over the Tibetan Plateau from TIPEX-III Cloud Radar Doppler Spectra. *Remote Sens.* **2017**, *9*, 964. [CrossRef]

33. Sokol, Z.; Minářová, J.; Novák, P. Classification of Hydrometeors Using Measurements of the Ka-Band Cloud Radar Installed at the Milešovka Mountain (Central Europe). *Remote Sens.* **2018**, *10*, 1674. [CrossRef]

34. Sokol, Z.; Zacharov, P.; Skripniková, K. Simulation of the storm on 15 August 2010, using a high resolution COSMO NWP model. *Atmos. Res.* **2014**, *137*, 100–111. [CrossRef]

35. Ge, J.; Zhu, Z.; Zheng, C.; Xie, H.; Zhou, T.; Huang, J.; Fu, Q. An improved hydrometeor detection method for millimeter-wavelength cloud radar. *Atmos. Chem. Phys.* **2017**, *17*, 9035–9047. [CrossRef]

36. BLIDS, Der Blitz Informationsdienst Von Siemens. Available online: https://new.siemens.com/global/de/produkte/services/blids.html (accessed on 10 March 2020).

37. Poelman, D.R.; Schulz, W.; Vergeiner, C. Performance Characteristics of Distinct Lightning Detection Networks Covering Belgium. *J. Atmos. Ocean. Technol.* **2013**, *30*, 942–951. [CrossRef]

38. Sokol, Z.; Mejsnar, J.; Pop, L.; Bližňák, V. Probabilistic precipitation nowcasting based on an extrapolation of radar reflectivity and an ensemble approach. *Atmos. Res.* **2017**, *194*, 245–257. [CrossRef]

39. Mejsnar, J.; Sokol, Z.; Minářová, J. Limits of precipitation nowcasting by extrapolation of radar reflectivity for warm season in Central Europe. *Atmos. Res.* **2018**, *213*, 288–301. [CrossRef]

40. Rakov, V.A.; Uman, M.A. *Lightning: Physics and Effects*; Cambridge University Press: Cambridge, UK, 2003; ISBN 978-0-521-58327-5.

41. Myagkov, A.; Seifert, P.; Wandinger, U.; Bauer-Pfundstein, M.; Matrosov, S.Y. Effects of Antenna Patterns on Cloud Radar Polarimetric Measurements. *J. Atmos. Ocean. Technol.* **2015**, *32*, 1813–1828. [CrossRef]

42. Kollias, P.; Albrecht, B.A.; Marks, F.D. Cloud radar observations of vertical drafts and microphysics in convective rain. *J. Geophys. Res. Atmos.* **2003**, *108*, 108. [CrossRef]

43. Lhermitte, R.M. Observation of rain at vertical incidence with a 94 GHz Doppler radar: An insight on Mie scattering. *Geophys. Res. Lett.* **1988**, *15*, 1125–1128. [CrossRef]

44. Rakov, V.A. *Fundamentals of Lightning*; Cambridge University Press: Cambridge, UK, 2016; ISBN 978-1-107-07223-7.

45. Minářová, J.; Sokol, Z.; Pešice, P. First comparison of measurements of Ka-band cloud radar with lightning in Central Europe. In Proceedings of the 2019 11th Asia-Pacific International Conference on Lightning (APL), Hong Kong, China, 12–14 June 2019; pp. 1–6.

MDPI

Article

Evaluation of GPM-Era Satellite Precipitation Products on the Southern Slopes of the Central Himalayas Against Rain Gauge Data

Shankar Sharma [1,2], **Yingying Chen** [1,3,*], **Xu Zhou** [1], **Kun Yang** [3,4], **Xin Li** [1,3], **Xiaolei Niu** [1], **Xin Hu** [1,2] and **Nitesh Khadka** [2,5]

[1] National Tibetan Plateau Data Center, Key Laboratory of Tibetan Environmental Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China; sharmash@itpcas.ac.cn (S.S.); xuzhou@itpcas.ac.cn (X.Z.); lixin@lzb.ac.cn (X.L.); xniu@itpcas.ac.cn (X.N.); huxin@itpcas.ac.cn (X.H.)

[2] University of Chinese Academy of Sciences, Beijing 100049, China; nkhadka@imde.ac.cn

[3] CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing 100101, China; yangk@itpcas.ac.cn

[4] Ministry of Education Key Laboratory for Earth System Modeling and Center for Earth System Science, Tsinghua University, Beijing 100084, China

[5] Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China

* Correspondence: chenyy@itpcas.ac.cn

Abstract: The Global Precipitation Measurement (GPM) mission provides high-resolution precipitation estimates globally. However, their accuracy needs to be accessed for algorithm enhancement and hydro-meteorological applications. This study applies data from 388 gauges in Nepal to evaluate the spatial-temporal patterns presented in recently-developed GPM-Era satellite-based precipitation (SBP) products, i.e., the Integrated Multi-satellite Retrievals for GPM (IMERG), satellite-only (IMERG-UC), the gauge-calibrated IMERG (IMERG-C), the Global Satellite Mapping of Precipitation (GSMaP), satellite-only (GSMaP-MVK), and the gauge-calibrated GSMaP (GSMaP-Gauge). The main results are as follows: (1) GSMaP-Gauge datasets is more reasonable to represent the observed spatial distribution of precipitation, followed by IMERG-UC, GSMaP-MVK, and IMERG-C. (2) The gauge-calibrated datasets are more consistent (in terms of relative root mean square error (RRMSE) and correlation coefficient (R)) than the satellite-only datasets in representing the seasonal dynamic range of precipitation. However, all four datasets can reproduce the seasonal cycle of precipitation, which is predominately governed by the monsoon system. (3) Although all four SBP products underestimate the monsoonal precipitation, the gauge-calibrated IMERG-C yields smaller mean bias than GSMaP-Gauge, while GSMaP-Gauge shows the smaller RRMSE and higher R-value; indicating IMERG-C is more reliable to estimate precipitation amount than GSMaP-Gauge, whereas GSMaP-Gauge presents more reasonable spatial distribution than IMERG-C. Only IMERG-C moderately reproduces the evident elevation-dependent pattern of precipitation revealed by gauge observations, i.e., gradually increasing with elevation up to 2000 m and then decreasing; while GSMaP-Gauge performs much better in representing the gauge observed spatial pattern than others. (4) The GSMaP-Gauge calibrated based on the daily gauge analysis is more consistent with detecting gauge observed precipitation events among the four datasets. The high-intensity related precipitation extremes (95th percentile) are more intense in regions with an elevation below 2500 m; all four SBP datasets have low accuracy (<30%) and mostly underestimated (by >40%) the frequency of extreme events at most of the stations across the country. This work represents the quantification of the new-generation SBP products on the southern slopes of the central Himalayas in Nepal.

Keywords: precipitation; GPM; IMERG; GSMaP; Nepal

1. Introduction

Precipitation is a vital component of the water cycle, and understanding the characteristics of precipitation is essential for hydro-meteorological applications [1,2]. In mountainous regions, water resource management is further challenging due to the complex climate associated with topographic variance [3]. In these regions, the occurrences of hydrological hazards such as floods, landslides and soil erosion are very sensitive to precipitation amounts. Thus, reliable and precise estimates of precipitation are a prerequisite for hydro-meteorological and natural disaster studies [4,5].

Nepal lies on the south-central part of the main Himalayan range, with more than 80% of the country covered by mountains; in this environment, there is a high probability of landslides and debris flows during the monsoon season. Precipitation in the country is extremely variable due to the complex topography. The seasonal cycle is predominantly governed by the monsoon system [6,7] with maximum (~80%) precipitation occurring in summer. Rain gauge-based measurements provide relatively accurate measurements of precipitation on the ground surface [8,9]. These observations developed by the Department of Hydrology and Meteorology (hereafter, DHM) in Nepal are relatively dense in the lowlands but sparse in high mountain areas [10,11]. The scarcity of rain gauge observations is a major challenge in hydro-meteorological studies and for effective water and disaster management. This scarcity of measurements also limits knowledge of precipitation patterns across the country [12]. Fortunately, high-resolution satellite-based precipitation (hereafter, SBP) products provide potential alternatives for monitoring precipitation on regular high-resolution grids, yielding unprecedented levels of detail especially over remote areas and mountainous regions where stations are very sparse. However, these estimates are indirect measurements and must be verified and calibrated using gauge observations before further application [13,14].

SBP estimates are based on various remotely sensed characteristics of clouds, such as cloud-top temperature (IR imagery), reflectivity (visible) or from the scattering effects of ice particles on passive microwave (PMW) radiation [15–18]. In the post- Tropical Rainfall Measuring Mission (TRMM) era, the Global Precipitation Measurement (GPM) Core Observatory spacecraft, equipped with advanced sensors and channels, like the Dual-frequency Precipitation Radar (DPR) and the GPM Microwave Imager (GMI) which had capabilities to sense light rain and snowfall, was launched on 27 February 2014 in a collaboration between NASA and the Japan Aerospace Exploration Agency (JAXA) [16]. New SBP products were introduced after the GPM mission: the Integrated Multi-satellite Retrievals for GPM (IMERG) [19,20]; meanwhile, JAXA updated to a newer version of the Global Satellite Mapping of Precipitation (GSMaP) product (GSMaP Version 07) with orographic rainfall correction [21].

Several studies have already evaluated SBP products around the globe for different hydro-meteorological applications [22–28]. They found that the new generation SBP products (GPM-Era) were improved than their previous version (TRMM-Era). For example, [29–31] found that the GPM-Era (IMERG-V3) precipitation product outperforms the TRMM-Era (TRMM 3B42V7, TMPA-RT) precipitation products. Meanwhile, studies conducted in Myanmar found the GSMaP-V07 product had the lowest accuracy when compared with the IMERG-05B product [32]. Similarly, Wang and Yong [33] also mentioned that IMERG-V05 performed better than GSMaP-V07, especially in high-elevation areas. A study conducted in Northwestern China did not found any significant difference in estimating the precipitation by IMERG-V06 and IMERG-V05 [34]. Nevertheless, several studies have concluded that new generation SBPs products can represent either the spatial pattern or the overall amount and general characteristics of extreme precipitation events over China [35,36], Cyprus [37], Austria [38], Africa [39] and India [40].

Besides several global studies, only a few studies have evaluated the SBPs in a topographically challenging region like Nepal. For example, the TRMM precipitation product shows negative bias (underestimation) as compared to gauge observations over the Himalayan region of the country [41]. Similarly, Islam et al. [12] found comparable results for 15 stations across the country. In contrast, Duncan and Biggs, [42] indicated that the TRMM product generally overestimated (positive bias) the precipitation as compared to a gauge-based gridded product (the Asian Precipitation

Highly Resolved Observational Data Integration Towards Evaluation: APHRODITE) over Nepal. In mountainous regions the TRMM (3B43) precipitation product shows reasonable skill, while the GSMaP, the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN), and the Climate Prediction Center Morphing Method (CMORPH) products showed considerably weaker performances in reproducing gauge-observed precipitation amounts [6]. A study in a high-elevation area (Khumbu Himalayas) of Nepal compared seasonal and diurnal variations of precipitation in TRMM (3B42), PERSIANN, CMORPH, and GSMaP products using hourly gauge observed precipitation [43]. They found that GSMaP performed poorly, while TRMM, PERSIANN and CMORPH had good agreement with rain-gauge data. Recently, Derin et al. [44] evaluated the GPM-era SBP products over different complex terrain areas, including ten stations from Nepal. They found that GSMaP-V07 was better for measuring the orographic precipitation and precipitation amount as compared to IMERG-V06B after the orographic rainfall classification ensemble in the GSMaP algorithm. The authors also noticed the better performance of IMERG-V05B to capture the light and heavy precipitation amount as compared to IMERG-V06B for the evaluated regions.

Most of the past studies in Nepal were based on previous-generation satellite products, which showed that errors in SBP estimates were partially related to the rugged topography as their algorithms could not detect orographically-induced precipitation appropriately. Additionally, the local climate and nature of the topography are some of the dominant factors to characterize the uncertainty of SBP products [35,45–47]. However, a systematic evaluation of the new-generation SBP products, and a intercomparison between these products, has not yet been performed at the national scale. Thus, in this study, we aimed to comprehensively evaluate four precipitation datasets from the two SBP products, i.e., GPM-era IMERG (V06B) and GSMaP (V07), against 388 gauge observations concerning their spatial and seasonal accuracy over Nepal. Their performances are analyzed for their tendencies and discrepancies depending on the different elevation range, and relative intensities on a daily and monthly timescale from March 2014 to December 2016. Moreover, the accuracy of East to West diversion of monsoon and extreme wet events on these SBP products are also analyzed. The result of this study will help to provide critical scientific references to choose the appropriate product for future scientific research.

2. Materials and Method

2.1. Study Area

Nepal is a country located on the southern slopes of the central Himalayas at 26.36°–30.45°N latitude and 80.06°–88.2°E longitude (Figure 1). Approximately 80% of the country comprises hills and mountains, and the remaining 20% is flatlands. The elevation of the country rises abruptly from the southern lowlands to the higher Himalayas giving rise to complex topography, weather and climate. Physiographically, the country is broadly classified into Terai (lowlands), hills and mountains [42,48]. The south Asian monsoon system and westerlies regulate the climatology of the country, with maximum precipitation in summer season (80%, June–September) followed by spring (March–May), autumn (October–November), and winter seasons (December–February) [49]. Usually, the monsoon advances from the east of the country in early June and promotes rainfall to the whole country within ten days [50]. During the winter, under the westerly-controlled climate in the western region, the country only receives about 3% of its annual precipitation [51]. Pre- and post-monsoon seasons are generally dry and hot, while the winter is cold and precipitation is generally in the form of snow, especially in high-elevation mountain areas [52].

2.2. Datasets

2.2.1. Rain Gauge Data

The daily precipitation dataset from 387 stations, between March 2014 and December 2016, were obtained from DHM Nepal (https://www.dhm.gov.np/contents/resources). The DHM stations provide daily datasets at 03 UTC. For the consistent measurement, DHM uses the same type of the United States standard eight-inch diameter manual rain gauges [53]. In addition to DHM datasets, data from a high-elevation Automatic Weather Station (AWS) located in the Everest region (27.95°N to 86.20°E, 5050 masl), Pyramid was also used.
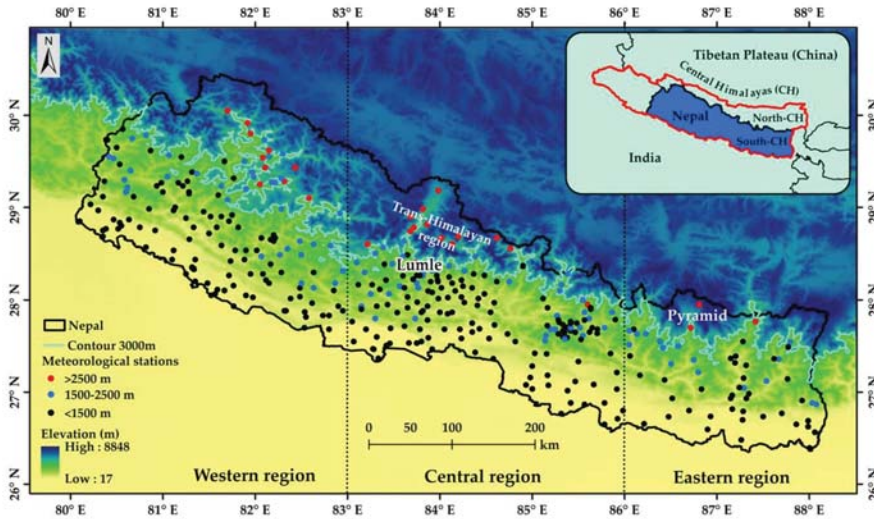
Figure 1. The study region, Nepal is divided into three different regions and the meteorological stations used are distributed from 60 m above sea level (masl) to 5050 masl. The extent of central Himalayas is adapted from Nie et al. [54].

In total, 388 stations' data were used for evaluation of SBP products (Figure 1), which were further subjected to quality control. The remote location, unavailability of AWS, lack of regular monitoring and maintenance of rain-gauge stations are the primary causes for discontinuities in the data series. Data coverage (%) at each station between March 2014 and December 2016 is presented in the Supplementary Materials (S1). The observations from 125 gauges were used in the development of the Global Precipitation Climatology Centre (GPCC) product, which was, in turn, used to calibrate IMERG precipitation totals [55]. About 54 stations also belonged to the Global Telecommunications System (GTS) and were used for the National Oceanic and Atmospheric Administration (NOAA)/CPC analysis, which was in turn used to calibrate the daily GSMaP-Gauge product [56]. Therefore, there is a potential dependency problem between the gauge observed and gauge-corrected SBP product used in this study. The mean precipitation (mm/day) at each rain-gauge station during the study period is presented in Figure 3a.

2.2.2. Satellite Datasets

IMERG is NASA's level-3 multi-satellite GPM product. After the GPM mission, four different versions (V03, V04, V05, and V06) of the IMERG algorithm were developed. IMERG combines retrievals from PMW and microwave-calibrated infrared (IR) to produce a high-resolution global SBP product. Compared to earlier versions (V03 and V04), IMERG V05, and V06 inter-calibrate all individual PMW

satellite estimates from Goddard PROFiling (GPROF2017) scheme [57]. Until V05B, the Lagrangian time interpolation scheme was computed from IR data, while in V06, a new model-based morphing scheme in which motion vectors are retrieved from Modern-Era Retrospective Reanalysis 2 (MERRA-2) and Goddard Earth Observing System Model (GEOS) Forward Processing (FP) data using total column water vapor (TQV) [58]. For the first time, IMERG V06 used Precipitation Retrieval and Profiling Scheme (PRPS) Sounder for Atmospheric Profiling of Humidity in the Intertropics by Radiometry (SAPHIR) only to the combined Ku-swath DPR/GMI product (CORRA) [20]. These intercalibrated estimates were then merged into a single PMW SBP estimate and used to calibrate the IR-based precipitation. IMERG uses a Kalman filter-based method to combine the observed PMW, propagated Pulse Width Modulation (PWM), and IR estimates into a single, best estimate. IMERG provides three different products, Early, Late and Final runs, with a latency of ~4 h, ~14 h, and ~3.5 months, respectively. The only forward direction of the cloud motion vector propagation algorithm was adopted in the Early run. In addition to Early, backward morphing was added in the Late run. Meanwhile, in the Final run, climatological calibration coefficients were added on the basis of Late run [33]. The rain-gauge data were not assimilated in Early and Late run, while the Final product was adjusted using 1° GPCC gauge analysis which is interpolated to 0.1° and applied equally to every half hour in the month. The Final run is mostly recommended for research purposes. IMERG applies the wind-loss correction scheme [59] to the GPCC gauge analyses during the calibration process. The latest IMERG version 06 products were released to the public in March 2019. The Final version of the IMERG product includes two precipitation fields, precipitationUncal and precipitationCal. The Early and Late precipitation products are identical to the Final precipitationUncal since the gauge correction is only applied to the precipitationCal field in the Final product. IMERG V6B had an upgrade of full intercalibration to GPM combined instrument datasets (2BCMB), also in this version the input precipitation rates were increased from 50 to 200 mm/h to adjust fractional coverage [44]. For this study, data from precipitationUncal (hereafter IMERG-UC) and precipitationCal (hereafter IMERG-C) of 0.1° spatial resolution from the Final run IMERG version 06B between March 2014 and December 2016 were obtained from the PMM website (https://gpm.nasa.gov/data-access/downloads/gpm). A detailed description of the IMERG algorithm can be found in Huffman et al. [60].

GSMaP is a SBP product developed by Japan Science and Technology (JST) under the Core Research for Evolutional Science and Technology (CREST) program [61,62]. To provide high-precision precipitation products, GSMaP combines various available PMW and IR sensors [63]. In the development of the GSMaP precipitation products, the instantaneous precipitation rate is first archived from the PMW radiometers based on various satellite platforms, such as GMI, advanced microwave sounding unit-A (AMSU-A), advanced microwave scanning radiometer-2 (AMSR-2), TRMM Microwave Imager (TMI), microwave humidity sounder (MHS), and special sensor microwave imager/sounder (SSMIS). Further, the gaps between the PMW-based estimates are propagated using atmospheric moving vector (cloud motion vector) calculated from successive IR images. In addition, a new Kalman filter model is applied to refine precipitation rates after the propagation [61]. Finally, forward and backward propagated precipitation estimates are weighted and combined to produce the GSMaP-MVK product. GSMaP-MVK also uses IR to correct satellite estimates but adopts different PMW imagers and sounders. In addition to PMW and IR, GSMaP-MVK estimates are adjusted using daily 0.5° NOAA/CPC gauge-based analysis to develop GSMaP-Gauge precipitation product; this reduces precipitation biases and has a latency of 3 days [64]. In addition to the GPM/DPR database, orographic rain correction classification also introduced in the algorithm of GSMaP-V07. In the current study, the Version 07 satellite-only (GSMaP-MVK) and gauge adjusted (GSMaP-Gauge) hourly datasets with a 0.1° spatial resolution were used. An overview of the selected datasets in the current study is provided in Table 1.

Table 1. Details of the datasets used in this study.

Datasets	Temporal Range	Spatial Resolution	Period	Coverage	Corrected by Gauges
DHM	1 day	388 stations	2014–2016		
IMERG-UC (V06B)	30 min	$0.1° \times 0.1°$	2014–2016	90°N–90°S	–
IMERG-C (V06B)	30 min	$0.1° \times 0.1°$	2014–2016	90°N–90°S	GPCC monthly
GSMaP-MVK (V07)	1 h	$0.1° \times 0.1°$	2014–2016	60°N–60°S	–
GSMaP-Gauge (V07)	1 h	$0.1° \times 0.1°$	2014–2016	60°N–60°S	CPC-daily

2.3. Methodology

2.3.1. Quality Control

All four SBP products are gridded datasets, and the gauge observed datasets are at the point scale, i.e., fixed at a single location on the ground surface. Therefore, a point-to-pixel comparison was performed to compare the point-based gauge observed data with the gridded precipitation datasets [65–68]. We extracted SBP estimates for the station locations instead of interpolating the gauge observations to avoid accumulating additional errors by gridding the observed data [69–71]. These SBP rates were first aggregated to obtain daily timescale records using DHM daily precipitation measurement time windows (03 UTC). Some of the station data feature missing values, and quality control was conducted for data consistency; if the gauge-observed daily data contained missing values, then the corresponding daily SBP data were simultaneously considered to be a missing value. The monthly data were computed when the station had more than 25 days of precipitation data available in a month; otherwise, the precipitation in that month was considered as a missing value. Similarly, the monthly SBP data were also considered a missing value for consistency if the corresponding monthly data were missing from the gauge observed datasets.

Mean, and summer (JJAS) mean monsoonal precipitation of SBP and gauge observed datasets during the study period were calculated for each station. Mean monthly regional datasets were computed for three regions, with stations located at longitudes of 80–82°E, 83–85°E and 86–88°E being grouped together as the western, central, and eastern regions, respectively (Figure 1). The stations were divided into three different physiographic regions (elevation intervals) to quantify the spatial patterns of SBP products. All stations below 1500 m, between 1500 and 2500, and above 2500 m, were aggregated into the low-elevation, mid-elevation, and high-elevation regions, respectively.

We also classified the gauge observed summer monsoon precipitation based on precipitation rate; all stations with mean monsoonal precipitation rates of less than 10 mm/day, between 10 and 20 mm/day, and higher than 20 mm/day was assigned to be lower, moderate, and higher precipitation zones, respectively. This allows analysis of SBP performances for different precipitation rates. The high-intensity related extreme rainfall events in four SBP datasets exceeding the 95th percentile (R95p) of observed precipitation was also examined. To do this, only those stations were selected when daily observed data was available more than 90% per year. The overall processes followed in the study is shown in Figure 2.

2.3.2. Statistical Analysis

Several statistical metrics were calculated to quantify the accuracy or differences between observation and estimated precipitation from SBP products based on monthly scale data. The correlation coefficient (R, Equation (1)) was used to measure the strength and direction of the linear association between datasets. The Relative Root Mean square Error (RRMSE, Equation (2)) which reflects the average magnitude of the deviation that a dataset will have from the gauge observed data; and mean bias (MB, Equation (3)) and absolute relative error (RE, Equation (4)) which measure any persistent tendency of a dataset to either overestimate or underestimate and the discrepancies between the magnitude of the estimated precipitation and the gauge observed dataset. Graphical

plots and different statistical measures were used to facilitate the inter-comparison between the SBP and gauge observed datasets. The formulae for the statistical metrics are:

$$R = \frac{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)\left(E_i - \overline{E}\right)}{\sqrt{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2}\sqrt{\sum_{i=1}^{n}\left(E_i - \overline{E}\right)^2}} \tag{1}$$

$$RRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(E_i - O_i)^2}{n}}}{\overline{O}} \tag{2}$$

$$MB = \frac{\left(\overline{E} - \overline{O}\right)}{n} \tag{3}$$

$$RE = \left|\frac{MB}{\overline{O}}\right| \tag{4}$$

where O is the gauge observation data, E is the estimated precipitation data using SBP products, $\overline{O}$ and $\overline{E}$ denotes the average value of their respective datasets, and n is the sample size. The perfect score for RRMSE, MB, and RE is ~0, while for R is 1.
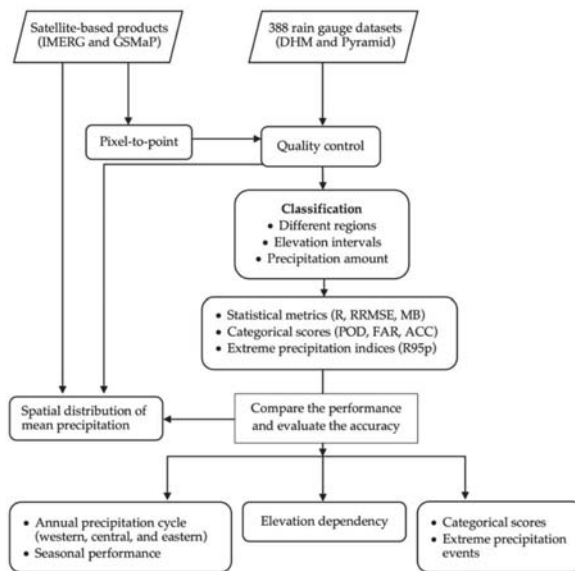
Figure 2. The flowchart of the overall processes followed in the study.

Additionally, a daily performance assessment was calculated for all SBP data based on categorical statistics. These statistics were computed for the individual stations to quantify the capacity to detect daily precipitation events. The statistics are based on a contingency table (Table 2) with two possible cases: a day with or without precipitation. In Table 2, a and d indicate the total events above 1 mm/day recorded by both datasets (gauge observed and SPB), while c and d indicate the total events recorded by both datasets below this threshold. This threshold value was selected to avoid the measurement error from the manual gauge system for the light precipitation amount (less than 1 mm/day).

Table 2. Contingency table to define daily precipitation based categorical scores for the evaluation of SBP with gauge observation.

		Gauge Observation	
		Precipitation	No-precipitation
SBP	Precipitation	a	b
	No-precipitation	c	d

In this study, three categorical indices were used for the assessment: the probability of detection (POD, Equation (5))—SBP's capacity to forecast the precipitation events accurately, and ranges from 0 to 1 (with 1 being an accurate score); False Alarm Ratio (FAR, Equation (6)), which represents how often the SBP's falsely detect a gauge observed precipitation event and ranges from 0 to 1 (with 0 being a perfect score); and Accuracy (ACC, Equation (7)), which is the fraction of all SBP product-based events that were correct, this has values ranging from 0 to 1, with 1 being a perfect score. All these metrics were computed for individual stations using respective daily precipitation series for the study period, with a threshold value of 1 mm/day to separate precipitation and no-precipitation events. The formulas for these statistical metrics are as follows:

$$POD = \frac{a}{(a + c)} \tag{5}$$

$$FAR = \frac{b}{(a + b)} \tag{6}$$

$$ACC = \frac{a + d}{(a + b + c + d)} \tag{7}$$

3. Results

3.1. Spatiotemporal Variability

3.1.1. The Spatial Pattern of Precipitation in Nepal

The spatial distributions of daily mean precipitation (mm/day) in observations and four different SBP datasets during the study period are presented in Figure 3. The observed datasets show large spatial variability of precipitation across the country. The highest mean precipitation amount (>10 mm/day) was observed in the Lumle areas (28.3°N, 84°E), whereas the low amount (<2 mm/day) in the high-elevation areas of central and western region (Figure 3a). Since the lowest precipitation area is located in the high-elevation areas of the central region, the high mountains remarkably block the atmospheric moisture from moving northward and considerably increase (decrease) precipitation in the southern (northern) slope of the central region. In the comparison of observed spatial distribution with the SBP datasets, all four SBP datasets generally showed the main characteristic, in which the high precipitation occurs in central Nepal. However, they differed largely in precipitation totals and location accuracy. The mean precipitation distribution from GSMaP-Gauge shows very similar characteristics, with the maximum precipitation (approximately 10–12 mm/day) at 28.3°N, 84°E (Figure 3e), whereas GSMaP-MVK shows the maximum precipitation (approximately 5–7 mm/day) at 28.5°N, 84°E (Figure 3c). In contrast, the IMERG-UC shows the maximum precipitation (approximately 10–12 mm/day) at 27.9°N, 84.8°E (Figure 3b), while IMERG-C shows high precipitation (approximately 4–5 mm/day) at 28.2° N, 84°E (Figure 3d). Another area (26.5°N, 88°E) of the highest rainfall in IMERG-C might be associated with the monsoon trough as seen over the lower ranches of the eastern region (Figure 3d). Notably, all four datasets are drier (<2 mm/day) in the high-elevation areas of central and western region (Figure 3b–e). The large scale patterns of the precipitation such as heavier orographic precipitation along with the southern slope of mountain ranges in the central region and lower precipitation (<2 mm/day) over the northern slope of central and western region (rain-shadow areas) is also qualitatively captured

by all four satellite precipitation datasets. It is worthy to note that, IMERG-UC showed better agreement with the spatial pattern and amount of precipitation in the observations than the IMERG-C, whereas GSMaP-Gauge showed significant improvement over the GSMaP-MVK. Moreover, all four SBP datasets tend to underestimate the mean precipitation across the country. GSMaP-Gauge (Figure 3e) well reproduces the overall spatial pattern of mean precipitation followed by IMERG-UC, GSMaP-MVK, and IMERG-C, respectively. The results suggest that gauge correction scheme for IMERG product requires further improvement in the study area.
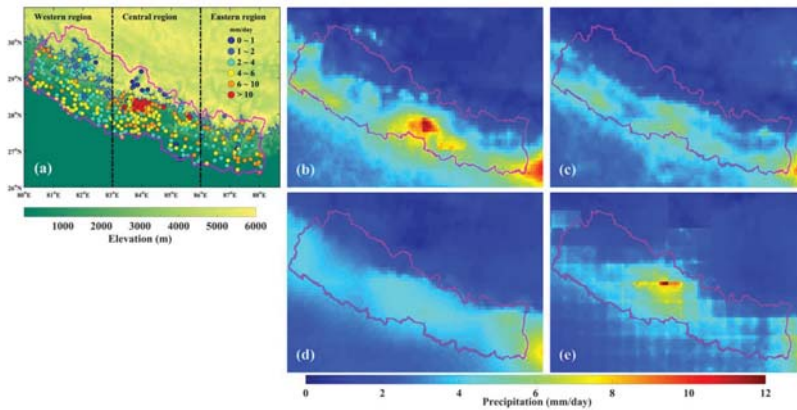
Figure 3. Spatial distribution of mean precipitation (mm/day) estimated by (**a**) observations, (**b**) IMERG-UC, (**c**) GSMaP-MVK, (**d**) IMERG-C and (**e**) GSMaP-Gauge averaged over March 2014 to December 2016. The magenta and black dotted lines represent the national boundary of the country and the three subregions, respectively.

3.1.2. Seasonal Pattern of Precipitation in Nepal

The annual cycles of precipitation in the gauge observed data, and all four SBP estimates in the western, central, and eastern regions of Nepal from 2015 to 2016 (only a two-year was chosen due to incomplete annual SBP datasets in 2014) are shown in Figure 4. High precipitation occurs from June–September in all three regions (Figure 4).
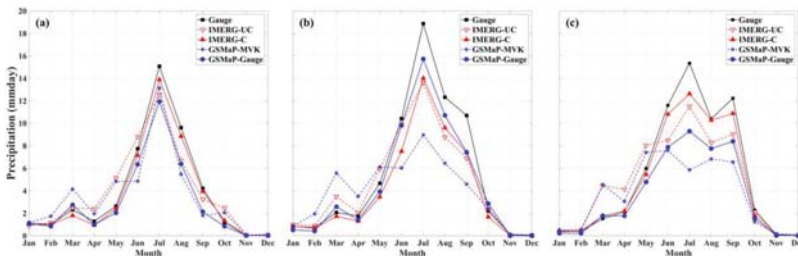
Figure 4. Monthly variation in precipitation (mm/day) over the (**a**) western (80–82ºE), (**b**) central (83–85ºE), and (**c**) eastern (86–88ºE) regions, derived from the gauge observed data and four SBP products averaged over 2015 to 2016.

The mean precipitation in the winter (December to February) was heavier over the western (0.74 mm/day) region than in the central (0.54 mm/day), and eastern regions (0.28 mm/day). In contrast, western Nepal was drier (4.90 mm/day) than the central (7.02 mm/day) and eastern regions (6.85 mm/day) during the other seasons (March to November) due to the influence of

summer monsoon. The precipitation in winter is primarily influenced by the westerlies system and is more pronounced in the western part of the country, while, moisture transfer from Bay of Bengal (monsoon) produces the widespread precipitation during the monsoon season (JJAS) over the country. All four SBP datasets show higher precipitation during the summer monsoon and lower precipitation in winter, with the maximum in July except for GSMaP-MVK in the eastern region (Figure 4c). The satellite-only datasets overestimated the precipitation during winter and pre-monsoon season; however, after the gauge calibration, the positive bias was reduced and is more consistent with observed datasets. Figure 4 indicates that among all four SBP datasets, the gauge calibrated datasets (i.e., IMERG-C and GSMaP-Gauge) represent well the seasonal precipitation variation across all three regions of Nepal, although they all yield underestimations. However, all four SBP datasets well captured the seasonal precipitation dynamics across the country.

For a detailed analysis, the statistical metrics of the four SBP datasets from 2015 to 2016 were calculated against the station observations (Table 3). In the western region, IMERG-UC and GSMaP-MVK showed smaller MBs than their gauge-calibrated datasets, i.e., IMERG-C and GSMaP-Gauge, respectively. Nevertheless, both gauge-calibrated datasets showed better overall performance as indicated by lower RRMSE and higher R-value (Table 3). For the central region, IMERG-UC showed the smaller MBs of −0.93 mm/day than that of −1.48 mm/day in IMERG-C; however, both have proximal RRMSE. Meanwhile, among all SBP, GSMaP-Gauge outperformed GSMaP-MVK and both IMERG datasets as indicated by the lowest MBs and RRMSE (Table 3). Both gauge-calibrated datasets showed very similar MBs with their corresponding satellite-only in the eastern region, although gauge calibrated IMERG-C performed more consistently, with a smaller RRMSE of 0.18, followed by GSMaP-Gauge. In the whole country, among all products, IMERG-UC showed the smallest MB of −0.47 mm/day and IMERG-C showed the lowest RRMSE of 0.28. It is worth noting that the positive bias in IMERG-UC and GSMaP-MVK between January and June later reduces the negative bias during July to October and shows smaller MBs among the datasets (Figures 4a–c and A1). The seasonal performances of all four SBP datasets were calculated to check the consistency in different seasons and presented in Table A1. The seasonal performance also showed that gauge calibrated datasets well represent the seasonal dynamics than satellite-only as indicated by lower MBs, RRMSE and higher R-value in Table A1. In general, all four datasets generally exhibited high correlations (R > 0.80), which indicate that the seasonal precipitation dynamics can be captured across the country by all four datasets.

Table 3. Statistical metrics in the western, central, and eastern regions, as well as in the whole study region, derived from the regional monthly mean precipitation (mm/day) from 2015 to 2016. Bold font indicates the best performance for a given metric.

Regions	Datasets	Mean (mm/day)	MB (mm/day)	RRMSE	R
Western region	IMERG-UC	3.86	**−0.03**	0.50	0.92
	IMERG-C	3.52	−0.31	**0.14**	**1.00**
	GSMaP-MVK	3.51	−0.32	0.82	0.81
	GSMaP-Gauge	2.97	−0.86	0.40	0.98
Central region	IMERG-UC	4.57	−0.93	0.48	0.95
	IMERG-C	4.02	−1.48	0.43	**1.00**
	GSMaP-MVK	3.92	−2.58	0.84	0.80
	GSMaP-Gauge	4.69	**−0.82**	**0.33**	0.98
Eastern region	IMERG-UC	4.75	−0.47	0.41	0.94
	IMERG-C	4.73	**−0.46**	**0.18**	**1.00**
	GSMaP-MVK	3.71	−1.50	0.71	0.83
	GSMaP-Gauge	3.64	−1.56	0.48	0.99
Whole country	IMERG-UC	4.31	**−0.47**	0.44	0.95
	IMERG-C	3.93	−0.86	**0.28**	**1.00**
	GSMaP-MVK	3.73	−1.05	0.76	0.81
	GSMaP-Gauge	3.86	−0.92	0.36	0.99

3.1.3. Elevation Dependency

The knowledge of the elevation gradient of precipitation is vital for many hydro-meteorological applications. As known that a larger portion of the precipitation occurs during the summer monsoon season, thus the elevation dependency was investigated based on monthly data from the summer monsoon season. The mean precipitation data from observed and four SBP were averaged over summer monsoon at different elevation ranges in every 500 m from 60 m to below 3000 m during the study period (Figure 5). The number of stations above 3000 m is very limited; thus, the elevation dependency was only calculated below 3000 m. The gauge observations reveal an evident elevation dependency of precipitation, as shown in Figure 5 for the monsoon period. Gauge observations show that precipitation gradually increases with increasing elevation up to 2000 m, and then decreases rapidly (black line, Figure 5a). The highest precipitation (approximately 13 mm/day) occurs in the range 1500–2000 m during the summer monsoon. These patterns are similar to the results revealed by the previous study conducted using gauge observations [72]. IMERG-C moderately captured this evident elevation-dependent pattern, with the highest precipitation (approximately 10 mm/day) in the elevation range of 1500–2000 m. In contrast, other three SBP products failed to capture this pattern; IMERG-UC and GSMaP-MVK showed the highest precipitation at the lowest elevation (below 500 m), and GSMaP-Gauge shows the highest precipitation in the elevation range 500–1000 m. In the higher elevation areas (above 3000 m) with limited stations (14), GSMaP-Gauge and both IMERG datasets overestimated the observed precipitation (not shown in Figure 5). This could be associated with the complex terrain and orographic effect [73,74]. It is worth to note that, orographic rain corrected GSMaP-Gauge showed the variation in precipitation amount for different elevation intervals (i.e., precipitation increase and decrease pattern) [75]. Additionally, we calculated the elevation dependency of SBP datasets by averaging the precipitation across all grid boxes within different elevation ranges (Figure 5b). The numbers of stations and grid boxes in different elevation ranges are listed in Table 4. Grid-based elevation dependency of SBP showed a similar pattern to that of the point-pixel results, but with slightly different precipitation amounts. Overall, both gauge satellite-only IMERG-UC and GSMaP-MVK significantly underestimated the monsoonal precipitation amount. However, after gauge correction, the precipitation estimates of gauge calibrated datasets were more consistent with the gauge observation than the satellite-only datasets. Therefore, the procedure of calibrating SBP products with rain gauge data is the reason for their increased accuracy.

Table 5 gives the statistical metrics of errors for the four SBP datasets across three different geographic regions, based on summer monsoon mean values at each station. In lowland areas (below 1500 m) the error metrics indicate that IMERG-UC showed smallest MBs of −0.85 mm/day; indicating the estimated precipitation amount was more consistent with the observed datasets; meanwhile, GSMaP-Gauge showed better overall performance than other three datasets with lower RRMSE (0.45) and a higher R-value (0.56). In the highest precipitating mid-elevation areas (between 1500 and 2500 m), both gauges calibrated datasets showed a more consistent performance to observed datasets than satellite-only datasets. IMERG-C showed the best performance to estimate the precipitation amount with lowest MBs of −3.07 mm/day, while GSMaP-Gauge presented the evident lowest RRMSE and higher R-value, indicating better performance in reproducing the spatial distribution of gauge observed precipitation among all. However, all four datasets underestimated the monsoon precipitation amount below 2500 m.
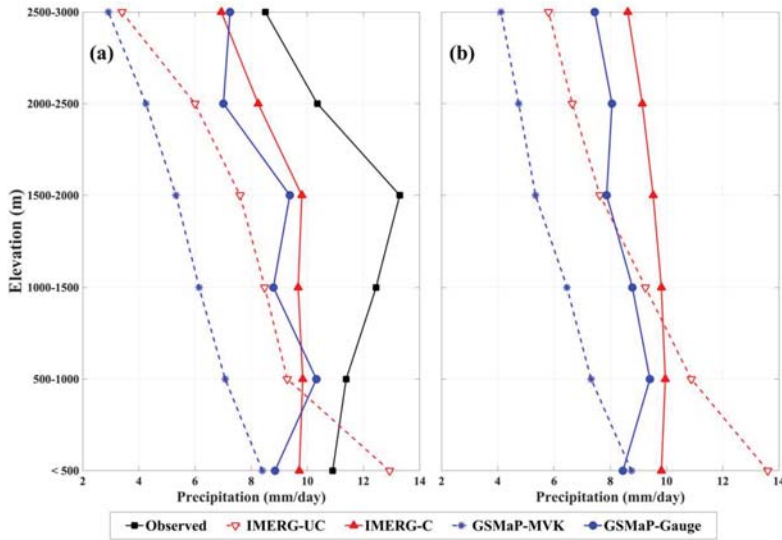
Figure 5. Mean precipitation averaged over the summer monsoon season (mm/day) for the study period, (**a**) in gauge observed and four SBP datasets at station locations, (**b**) in IMERG and GSMaP product at each grid box averaged over different elevation ranges, respectively.

Table 4. Numbers of stations and grid boxes in different elevation ranges.

Elevation Range	No. of Stations	No. of Grids
Below 500 m	121	354
500–1000	56	173
1000–1500	90	185
1500–2000	67	126
2000–2500	27	98
2500–3000	13	79

Table 5. Statistical metrics in the station mean precipitation at different elevation intervals (below 1500 m, between 1500 and 2500 m and above 2500 m), as well as in the whole study region, derived from the four SBP datasets and compared to the gauge observations during the summer monsoon for the study period. Bold font indicates the dataset with the best performance for a given metric.

Regions	Datasets	Mean (mm/day)	MB (mm/day)	RRMSE	R	RE (%)
Below 1500 m (267 stations)	IMERG-UC	10.67	**−0.85**	0.57	0.08	**7.36**
	IMERG-C	9.73	−1.79	0.47	0.41	15.58
	GSMaP-MVK	7.36	−4.16	0.58	0.36	36.08
	GSMaP-Gauge	9.14	−2.38	**0.45**	**0.56**	20.66
Between 1500 and 2500 m (95 stations)	IMERG-UC	7.11	−5.31	0.70	0.24	42.75
	IMERG-C	9.34	**−3.07**	0.56	0.46	**24.74**
	GSMaP-MVK	4.99	−7.42	0.78	0.49	59.82
	GSMaP-Gauge	8.72	−3.69	**0.53**	**0.63**	29.72
Above 2500 m (26 stations)	IMERG-UC	3.10	−2.18	1.15	0.42	41.28
	IMERG-C	6.00	**0.72**	**1.07**	**0.65**	**16.63**
	GSMaP–MVK	2.32	−2.96	1.16	0.59	56.07
	GSMaP–Gauge	7.78	1.50	1.44	−0.02	28.48
Whole country (388 stations)	IMERG-UC	9.29	−2.03	0.63	0.19	18.13
	IMERG-C	9.38	**−1.94**	0.52	0.48	**17.19**
	GSMaP-MVK	6.44	−4.88	0.66	0.40	43.10
	GSMaP-Gauge	8.88	−2.44	**0.51**	**0.55**	21.56

In high-elevation regions (above 2500 m), characterized by complex topography with low precipitation, IMERG-C showed the best performance with smaller errors (MB and RRMSE) and higher R than other datasets, demonstrating that the calibration based on GPCC data significantly improved the IMERG product. Meanwhile, the GSMaP-Gauge product overestimated summer monsoon precipitation and showed very poor correlation with the gauge observation. As mentioned in Section 2.2.1, the GPCC data merged observations from 125 gauges in Nepal, while NOAA/CPC data only merged observations from 54 gauges. Therefore, GPCC data may integrate more precipitation information, especially in high elevation regions where gauge observations are very scarce than NOAA/CPC data. This might be the reason for the improved performance of IMERG-C than GSMaP-Gauge in high-elevation regions.

Overall, the gauge-calibrated products performed better than the satellite-only products on a monthly scale. IMERG-C yielded smaller MBs than GSMaP-Gauge, while GSMaP-Gauge showed the smaller RRMSE and higher R-value; indicating IMERG-C was more consistent to estimate the precipitation amount than GSMaP-Gauge, whereas GSMaP-Gauge presented more reasonable spatial distribution than IMERG-C.

3.2. Performance-Based on Daily Data

3.2.1. Statistical Scores

Figure 6 illustrates the spatial distributions of the POD, FAR and ACC values in the four SBP datasets at each station across the country. These values were calculated based on daily precipitation data. POD is above 70% at most of the stations in GSMaP-Gauge, followed by 40–80% in IMERG product, and 40–60% in GSMaP-MVK. Therefore, the daily gauge-calibrated GSMaP-Gauge outperformed the other three datasets to detect gauge precipitation events. It is worth noting that POD is higher in GSMaP-Gauge compared to GSMaP-MVK, while similar performances are found in both IMERG datasets. The GSMaP-Gauge datasets were calibrated based on daily scale NOAA/CPC data; therefore, both the amount and the occurrence were corrected, while IMERG-C datasets were corrected based on monthly scale GPCC data, thus only the precipitation amounts were adjusted. This may be the reason for similar POD performance in two IMERG products. All four datasets showed similar FAR distributions, with the best performance in mid-elevation areas and poor performance in high-elevation areas, revealing that the error was lower when the precipitation amount was higher. ACC exceeded 70% in all four SBP datasets at most of the selected stations across the country (Figure 6).

POD, FAR, and ACC were also calculated for the three different elevation intervals (Table 6). The daily gauge-calibrated GSMaP-Gauge performed fair well at detecting the gauge observed precipitation events with PODs of 0.73, 0.71 and 0.66 for elevations below 1500 m, 1500–2500 m, and above 2500 m, respectively. Notably, differences in the ACC and FAR scores were nominal (Table 6). In general, all four datasets performed with acceptable scores (ACC higher than 0.70) in detecting both the precipitation and no-precipitation events across the country.
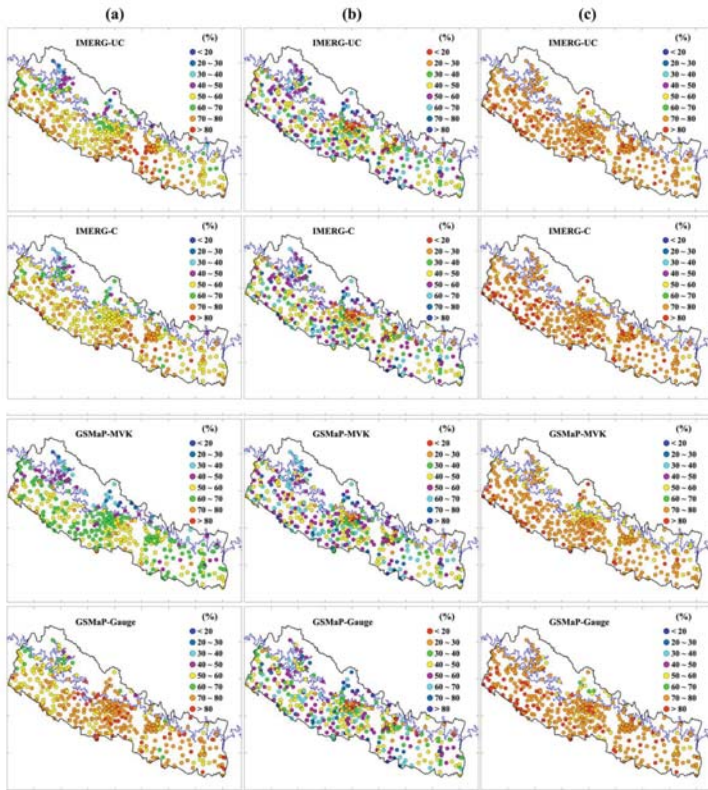
Figure 6. SBP performance at individual station measurements expressed in the form of the panel (**a**) POD, panel (**b**) FAR, and panel (**c**) ACC in % at each station during the study period. The black and blue lines show the national boundary and 3000 m elevation contour, respectively.

Table 6. Performances of the four SBP datasets expressed by POD, FAR and ACC at each station, averaged over three elevation intervals (<1500 m, 1500–2500 m and > 2500 m), as well as for the whole study region during the study period. A threshold value of 1 mm/day was selected to separate precipitation and no-precipitation events. Bold font indicates the dataset with the best performance for a given metric.

Regions	Datasets	POD	FAR	ACC
Below 1500 m	IMERG-UC	0.66	0.49	0.76
	IMERG-C	0.69	**0.48**	**0.77**
	GSMaP-MVK	0.60	0.52	0.75
	GSMaP-Gauge	**0.73**	0.50	**0.77**
Between 1500 and 2500 m	IMERG-UC	0.59	0.42	0.74
	IMERG-C	0.61	**0.41**	0.75
	GSMaP-MVK	0.54	0.47	0.71
	GSMaP-Gauge	**0.71**	0.41	**0.76**
Above 2500 m	IMERG-UC	0.49	0.55	**0.73**
	IMERG-C	0.54	**0.54**	0.72
	GSMaP-MVK	0.37	0.62	0.72
	GSMaP-Gauge	**0.66**	0.57	0.70
Whole country	IMERG-UC	0.62	0.48	0.75
	IMERG-C	0.64	**0.47**	**0.76**
	GSMaP-MVK	0.57	0.52	0.74
	GSMaP-Gauge	**0.72**	0.48	**0.76**

3.2.2. Extreme Precipitation Events

The spatial patterns of extreme precipitation events identified by the four SBP datasets are presented in Figure 7.
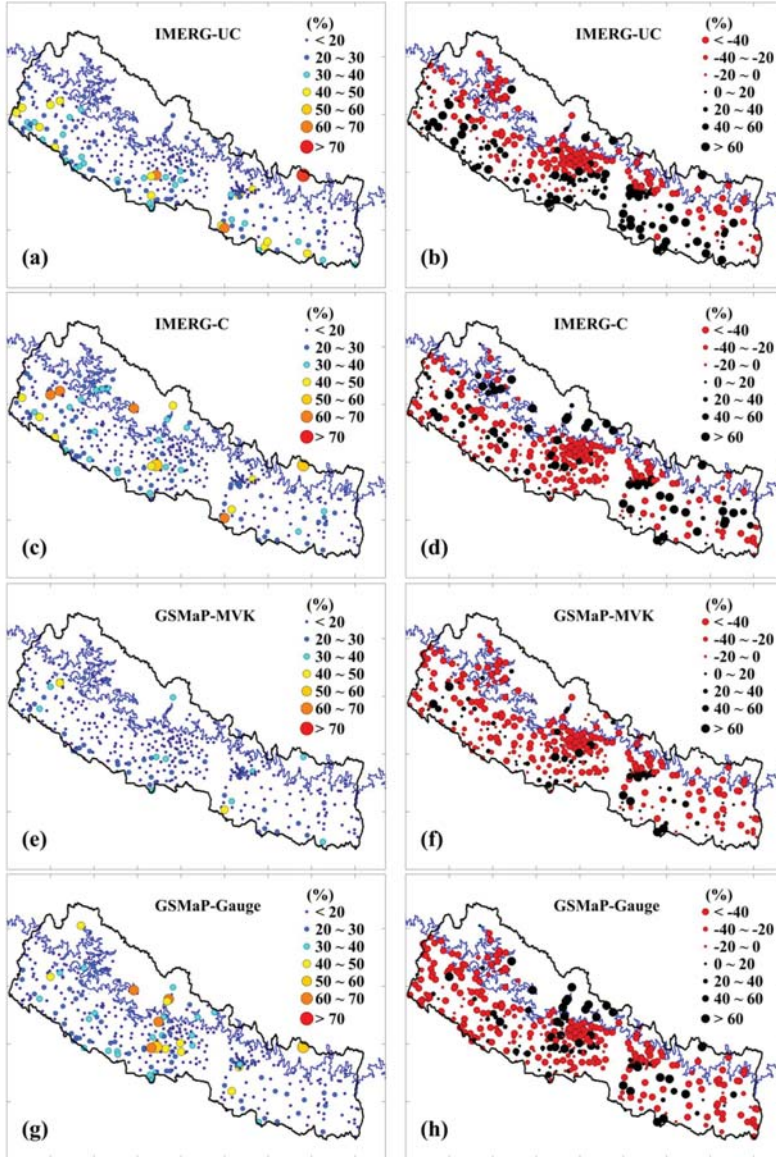
Figure 7. Spatial distribution of the daily precipitation 95th percentile from 2014 to 2015 in the four SBP datasets. Figure (**a**,**c**,**e**,**g**) shows the accurately detected extreme precipitation events in (%), (**b**,**d**,**f**,**h**) bias in extreme precipitation events for each station. The black and blue lines denotes the national boundary of the country and 3000 masl elevation contour, respectively.

Extreme precipitation events are defined as those exceeding the 95th percentile (high-intensity extreme) values in gauge-observed datasets from 2015 to 2016 for each station. Extreme events were calculated only for those stations with daily observation data available for more than 90% of the year. Figure 7 shows the extreme events detected by SBP datasets on the same day as those in the gauge-observed data sets (temporal accuracy) and the mean bias in the total number of extreme events at each station across the country, respectively. GSMaP-Gauge has been moderately improved in contrast with GSMaP-MVK, especially in central Nepal, where more precipitation was observed than in other areas (Figure 7e,g). The spatial distribution of the extreme events suggests that all four SBP datasets have low accuracy and mostly underestimate the frequency of extreme events over the study area (Figure 7b,d,f,h).

Figure 8 shows the performances of the four SBP datasets in detecting extreme events within the three elevation intervals. The statistics indicate that all four SBP datasets underestimated the frequency of extreme events in regions below 2500 m, while the IMERG-C and GSMaP-Gauge showed many more fake extreme events than the satellite-only products and thus overestimated the frequency in regions above 2500 m. As shown in Figure 8a, the higher number of extreme events were observed in regions below 2500 m (low and mid-elevation) than that for high elevation regions, and most DHM gauge stations (96.5% of total) are also located in these regions. Therefore, the GPCC analysis dataset interpolated from data of 125 DHM gauge stations and the NOAA/CPC dataset interpolated from data of 54 DHM gauge stations, which were used to calibrate the IMERG and GSMaP products respectively, may present fake high occurrence of extreme events in regions above 2500 m. That is why the calibrated SBP products overestimated the frequency of extreme events in regions above 2500 m. According to Figure 8, IMERG-UC performed much better in presenting the extreme event occurrence than other products especially in regions below 1500 m, suggesting that sometimes the calibration may skew some important signals contained in the satellite-only product. In general, all four SBP products had low accuracy (Figure 8b) and underestimated the frequency of extreme events (Figure 8a) across the country.
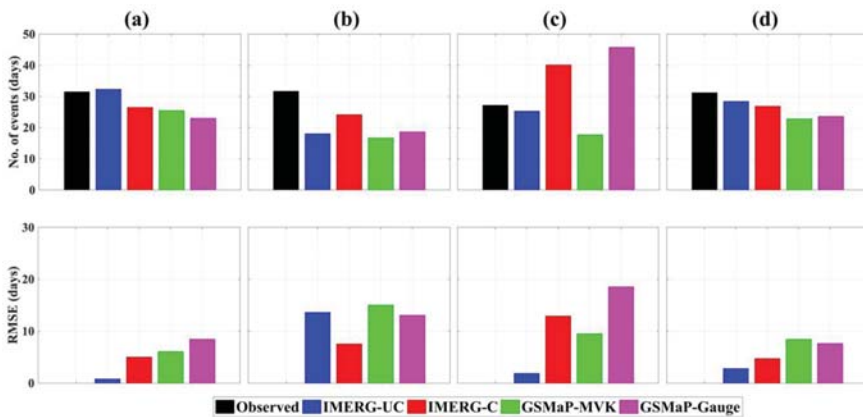
Figure 8. Bar charts showing extreme events (mean number of days) and RMSE (mean number of days) in the gauge observed data and four SBP datasets from 2015 to 2016 for (**a**) low-elevation (below 1500 m), (**b**) mid-elevation (between 1500 and 2500 m), (**c**) high-elevation (above 2500 m), and (**d**) the whole region.

4. Discussion

SBP products provide new alternatives for station observations; however, uncertainties are associated with both gauge observations and SBP estimations. Unfortunately, these uncertainties are difficult to quantify and may have influenced the above results.

Gauge instruments may suffer from systematic biases caused by wind-induced evaporation loss and underestimation of trace values. These biases are more prominent during the winter due to the lower precipitation totals and a higher prevalence of snow. Due to the lack of automatic gauge stations in high-elevation areas, conventional rain gauges measure qualitative precipitation amounts (rainfall + snowfall). Therefore, the evaluation of datasets during the summer is more reliable than during the winter, and the evaluation in this study mainly focuses on the summer monsoon season when the effect of evaporation loss in the observations is not significant due to the large precipitation totals. It is worth noting that the gauge-observed datasets used in this study are not wind corrected, due to the lack of wind speed data for the selected stations, but they were used to evaluate the IMERG-gauge datasets which are calibrated by the wind-loss corrected GPCC gauge analyses. This fact also weakened the certainty of the evaluation results in Section 3. Besides, among the DHM stations used in this study, data from 125 stations were merged to produce the GPCC gauge analysis, which was used to calibrate the GPM-IMERG product. Similarly, data from about 54 stations were also merged to produce the NOAA/CPC gauge-based analysis, which was used to correct the GSMaP product. The overlaps may lead to underestimation of the evaluation errors, which were not identified due to lack of information.

Furthermore, previous studies revealed a quite large variability of precipitation in the high-elevation areas of Nepal with the importance of nocturnal precipitation [43]. In these regions, most of DHM gauge stations are located in valley bottoms [76,77], where the nocturnal precipitation prevails. In such regions, the daily DHM gauge data do not capture the representative precipitation variability and may leave a gap in the performance quantification of SBP products. On the other hand, since the calibrated SBP products are corrected using the gauge-based analysis datasets, which present finer spatial patterns than the natural pattern, especially in mountainous regions. However, the calibrated SBP products may also smoothen the ture spatial pattern in mountainous regions, and even skew some important signals contained in satellite-only products.

Several previous studies also mentioned that topographic nature and regional climate are some dominant factors that influence the precipitation retrieval algorithm used in TRMM and GPM precipitation datasets [78]. The accuracy of the SBP precipitation data depends on various factors, such as regional effects and precipitation intensities. The scatter plot of precipitation rates between gauge observed, and SBP products were drawn using monthly precipitation data during the summer monsoon to quantify the performance of the SBP datasets for different precipitation intensities (Figure 9).

GSMaP-Gauge, IMERG-C, IMERG-UC and GSMaP-MVK performed better for low precipitation rates (<10 mm/day) than for high precipitation rates (>10 mm/day), indicating that the SBP datasets have difficulty in estimating heavy precipitation. These SBP datasets also overestimated the amount of light rain and underestimated the amount of heavy rain. Such underestimation of heavy rain could be the reason for underestimated high-intensity related gauge observed extreme events (Figures 7b and 8) across the country. These discrepancies are primarily related to false precipitation in the form of light rain or solid precipitation and underestimation of heavy precipitation, respectively. Also, complex physiographic nature of the study region may have an effect on the upward microwave radiation, which makes difficult for the satellite to resolve precipitation over areas with low precipitation amount, especially in the mountainous region [79,80].

SBP datasets are indirect measurements that are based on satellite/sensor constellations, including PMW and IR sensors onboard LEO and geostationary satellites. These datasets may not accurately detect precipitation in high-elevation areas [81], especially in the winter season, when the ground surface is covered with snow and ice [13]. The errors of precipitation estimated by PMW are mainly based on scattering signal which cannot catch up warm/low-level precipitation frequently occurs

in low-elevation areas and algorithm to interpolate finer time-scales, such as using cloud motion vector by IR. Meanwhile, the used algorithm can not capture to interpolate finer time-scales, such as using cloud motion vector by IR.
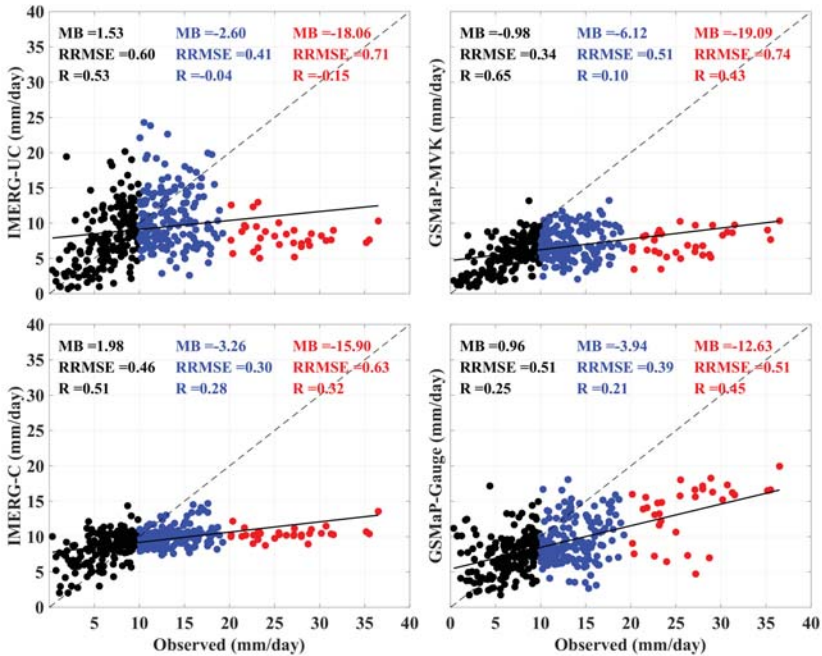
Figure 9. Scatter plot of differences in precipitation rates between gauge observed and SBP products, derived from monthly mean precipitation averaged over the monsoon season (mm/day). All the units of statistical metrics are in mm/day. Black, blue, and red colors indicate the performance statistics for precipitation rates less than 10 mm/day, between 10 and 20 mm/day, and above 20 mm/day, respectively. The continuous black and dotted black line represents the linear regression and 1:1 line, respectively.

To reduce such bias, satellite-only estimates were calibrated using gauge-based GPCC and CPC datasets. The performances of the IMERG-C and GSMaP-Gauge datasets were also influenced by the quality and temporal range of the calibrated gauge-based GPCC and CPC datasets, respectively. Meanwhile, satellite-only datasets are only effectively adjusted for those areas where gauge data are available. Our results showed a substantial improvement in gauge-calibrated SBP datasets, which are more consistent than satellite-only datasets, due to the advantages of observed gauge adjustments. This result is similar to studies conducted in Central Asia [82], China [64,83], East Africa [84] and Ethiopia [85]. However, deterioration of IMERG-C in low-elevation areas (Figures 3 and 7 and Table 5) as compared to IMERG-UC, as well as deterioration of GSMaP-Gauge in high-elevation areas (Figure 8c and Table 5) as compared to GSMaP-MVK, might be related to limitations of adjusted relevant rain-gauge density (GPCC and CPC). Such discrepancies indicate that the IMERG-C and GSMaP-Gauge retrieval algorithms need further improvements, particularly for mountainous areas, such as Nepal. Additionally, both PMW and IR satellites have complication in detecting shallow orographic precipitation [61,86,87]. We found that the local weather conditions and nature of the topography also influence the rainfall capturing capacity of SBP product.

5. Conclusions

This study attempted to evaluate the latest four SBP products in the southern slope of central Himalayas, Nepal and compare both satellite-only (IMERG-UC and GSMaP-MVK) and gauge-calibrated (IMERG-C and GSMaP-Gauge) products for their accuracy and discrepancies with 388 gauges measurements from March 2014 to December 2016. Conventional statistical metrics and categorical scores were used to quantify the performances of these SBP products.

Precipitation estimates differ widely between SBP products, depending on the season and location. The GSMaP-Gauge dataset was more consistent at representing the spatial pattern of observed precipitation followed by IMERG-UC, GSMaP-MVK, and IMERG-C. However, all four datasets can capture the seasonal precipitation dynamics across the country. Among them, IMERG-C and GSMaP-Gauge presented more consistent seasonal dynamics range (in terms of RRMSE and R) with the gauge observations than the satellite-only datasets. Even though all four SBP products underestimate the gauge-observed precipitation across Nepal; both gauge-calibrated SBP datasets performed better (lower RRMSE, higher R) than the satellite-only datasets. IMERG-C and GSMaP-Gauge showed similar errors (MB and RRMSE) in Nepal, although both had discrepancies in capturing the precipitation patterns. For instance, GSMaP-Gauge presented a more reasonable spatial distribution, while IMERG-C moderately reproduced the evident elevation-dependent pattern of precipitation as revealed by gauge observations, i.e., increasing precipitation with an increasing elevation below 2000 m and then decreasing above 2000 m.

When selecting 1 mm/day as the threshold defining a daily rainfall event, benefit from merging daily gauge-based NOAA/CPC analysis data, GSMaP-Gauge performed best (with higher POD) for detecting gauge observed precipitation events among four datasets. Gauge observations indicated that more high-intensity precipitation extreme events (95th percentile) occur in regions with an elevation below 2500 m. All four SBP datasets underestimated the total frequency of extreme precipitation events across the country. It is worth noting that IMERG-UC performed much better in presenting the occurrence of extreme events than other products especially in regions below 1500 m, suggesting that sometimes the calibration may skew some important signals contained in the satellite-only product.

The present work addresses the lack of systematic evaluation of the latest two SBP products in the southern slope of central Himalayas, Nepal. This evaluation provides a statistical basis and allows rigorous data selection in meteorological, hydrological, glaciological, and disaster-related studies within the study region. We recommend that further evaluation of SBP products based on the weather characteristics over the complex terrain may provide useful information to algorithm developers and data users.

Appendix A

Table A1. Statistical metrics in the different seasons, derived from the monthly mean precipitation (mm/day) from 2015 to 2016. Bold font indicates the best performance for a given metric.

Seasons	Datasets	Mean (mm/day)	MB (mm/day)	RRMSE	R
Pre-monsoon (March-May)	IMERG-UC	3.99	1.33	0.59	0.98
	IMERG-C	2.18	−0.48	**0.2**	**0.99**
	GSMaP-MVK	4.6	1.94	0.85	0.88
	GSMaP-Gauge	2.45	**−0.21**	0.22	0.93
Monsoon (June–September)	IMERG-UC	8.86	−2.58	0.29	0.9
	IMERG-C	9.32	**−2.11**	**0.2**	**1**
	GSMaP-MVK	6.5	−4.93	0.52	0.67
	GSMaP-Gauge	8.94	−2.5	0.25	0.95
Post-monsoon (October–November)	IMERG-UC	1.23	0.26	0.81	0.79
	IMERG-C	0.82	−0.15	0.22	**1**
	GSMaP-MVK	1.12	0.15	0.64	0.83
	GSMaP-Gauge	0.99	**0.02**	**0.08**	**1**
Winter (December–Feburary)	IMERG-UC	0.48	**0.02**	0.2	0.98
	IMERG-C	0.42	−0.03	**0.08**	**1**
	GSMaP-MVK	0.7	0.25	0.97	0.84
	GSMaP-Gauge	0.32	−0.13	0.35	0.94
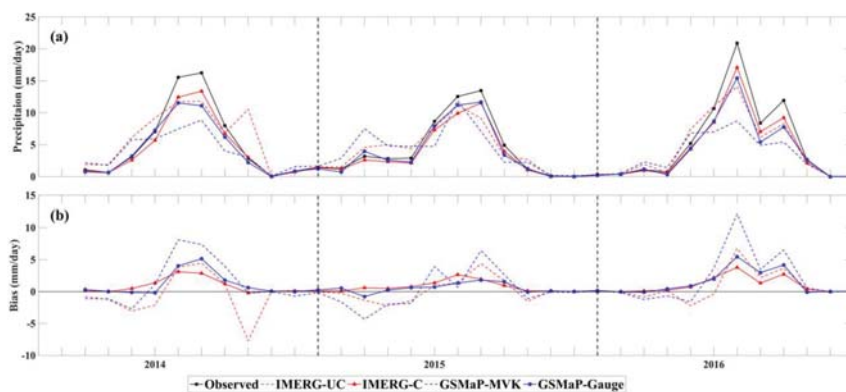
Figure A1. (**a**) The timeseries of monthly mean precipitation and (**b**) Bias between observed and four SBP datasets averaged over the study area.

References

1. Daly, C.; Slater, M.E.; Roberti, J.A.; Laseter, S.H.; Swift, L.W. High-resolution precipitation mapping in a mountainous watershed: Ground truth for evaluating uncertainty in a national precipitation dataset. *Int. J. Climatol.* **2017**, *37*, 124–137. [CrossRef]
2. Schneider, U.; Ziese, M.; Meyer-Christoffer, A.; Finger, P.; Rustemeier, E.; Becker, A. The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources. *P. Int. Ass. Hydrol. Sci.* **2016**, *374*, 29–34. [CrossRef]
3. Viviroli, D.; Weingartner, R. The hydrological significance of mountains: From regional to global scale. *Hydrol. Earth Syst. Sci.* **2004**, *8*, 1016–1029. [CrossRef]
4. Viviroli, D.; Durr, H.H.; Messerli, B.; Meybeck, M.; Weingartner, R. Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resour. Res.* **2007**, *43*. [CrossRef]
5. Viviroli, D.; Archer, D.R.; Buytaert, W.; Fowler, H.J.; Greenwood, G.B.; Hamlet, A.F.; Huang, Y.; Koboltschnig, G.; Litaor, M.I.; Lopez-Moreno, J.I.; et al. Climate change and mountain water resources: Overview and recommendations for research, management and policy. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 471–504. [CrossRef]
6. Krakauer, N.Y.; Pradhanag, S.M.; Lakhankar, T.; Jha, A.K. Evaluating Satellite Products for Precipitation Estimation in Mountain Regions: A Case Study for Nepal. *Remote Sens. (Basel)* **2013**, *5*, 4107–4123. [CrossRef]

7. Sharma, S.; Khadka, N.; Hamal, K.; Baniya, B.; Luintel, N.; Joshi, B.B. Spatial and Temporal Analysis of Precipitation and Its Extremities in Seven Provinces of Nepal (2001–2016). *Appl. Ecol. Environ. Sci.* **2020**, *8*, 64–73.

8. Sun, Q.H.; Miao, C.Y.; Duan, Q.Y.; Ashouri, H.; Sorooshian, S.; Hsu, K.L. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Rew. Geophys.* **2018**, *56*, 79–107. [CrossRef]

9. Petersen, W.A.; Christian, H.J.; Rutledge, S.A. TRMM observations of the global relationship between ice water content and lightning. *Geophys Res. Lett.* **2005**, *32*. [CrossRef]

10. Diodato, N.; Tartari, G.; Bellocchi, G. Geospatial Rainfall Modelling at Eastern Nepalese Highland from Ground Environmental Data. *Water Resour. Manag.* **2010**, *24*, 2703–2720. [CrossRef]

11. Barros, A.P.; Lang, T.J. Monitoring the monsoon in the Himalayas: Observations in central Nepal, June 2001. *Mon. Weather Rev.* **2003**, *131*, 1408–1427. [CrossRef]

12. Islam, M.; Das, S.; Uyeda, H. Calibration of TRMM derived rainfall over Nepal during 1998–2007. *Open Atmospheric Sci. J.* **2010**, *4*, 12–23. [CrossRef]

13. Tian, Y.D.; Peters-Lidard, C.D. A global map of uncertainties in satellite-based precipitation measurements. *Geophys. Res. Lett.* **2010**, *37*. [CrossRef]

14. Khairul, I.M.; Mastrantonas, N.; Rasmy, M.; Koike, T.; Takeuchi, K. Inter-Comparison of Gauge-Corrected Global Satellite Rainfall Estimates and Their Applicability for Effective Water Resource Management in a Transboundary River Basin: The Case of the Meghna River Basin (vol 6, 828, 2018). *Remote Sens.* **2018**, *10*, 828. [CrossRef]

15. Seto, S.; Iguchi, T.; Oki, T. The Basic Performance of a Precipitation Retrieval Algorithm for the Global Precipitation Measurement Mission's Single/Dual-Frequency Radar Measurements. *IEEE Trans. Geosci. Remote. Sens.* **2013**, *51*, 5239–5251. [CrossRef]

16. 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]

17. Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]

18. Tang, G.Q.; Ma, Y.Z.; Long, D.; Zhong, L.Z.; 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]

19. Huffman, G.J. Algorithm Theoretical Basis Document (ATBD) Version 06. NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG), NASA. Available online: https://pmm.nasa.gov/data-access/downloads/gpm (accessed on 10 January 2020).

20. Tan, J.; Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. IMERG V06: Changes to the Morphing Algorithm. *J. Atmospheric Ocean. Technol.* **2019**, *36*, 2471–2482. [CrossRef]

21. Kubota, T.; Aonashi, K.; Ushio, T.; Shige, S.; Takayabu, Y.N.; Arai, Y.; Tashima, T.; Kachi, M.; Oki, R. Recent progress in global satellite mapping of precipitation (GSMaP) product. Proceedings of 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 2712–2715.

22. Shen, Y.; Xiong, A.Y. Validation and comparison of a new gauge-based precipitation analysis over mainland China. *Int. J. Climatol.* **2016**, *36*, 252–265. [CrossRef]

23. Tan, J.; Petersen, W.A.; Kirstetter, P.E.; Tian, Y.D. Performance of IMERG as a Function of Spatiotemporal Scale. *J. Hydrometeorol.* **2017**, *18*, 307–319. [CrossRef]

24. Brown, J.E. 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]

25. Zhu, Q.; Luo, Y.; Zhou, D.; Xu, Y.-P.; Wang, G.; Gao, H. Drought Monitoring Utility using Satellite-Based Precipitation Products over the Xiang River Basin in China. *Remote Sens.* **2019**, *11*, 1483. [CrossRef]

26. Satgé, F.; Ruelland, D.; Bonnet, M.-P.; Molina, J.; Pillco, R. Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 595–619. [CrossRef]

27. 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]

28. Tan, X.; Yong, B.; Ren, L. Error features of the hourly GSMaP multi-satellite precipitation estimates over nine major basins of China. *Hydrol. Res.* **2017**, *49*, 761–779. [CrossRef]

29. Yang, X.; Lu, Y.; Tan, M.L.; Li, X.; Wang, G.; He, R. Nine-Year Systematic Evaluation of the GPM and TRMM Precipitation Products in the Shuaishui River Basin in East-Central China. *Remote Sens.* **2020**, *12*, 1042. [CrossRef]

30. Gebregiorgis, A.S.; Kirstetter, P.E.; Hong, Y.E.; Gourley, J.J.; Huffman, G.J.; Petersen, W.A.; Xue, X.; Schwaller, M.R. To what extent is the day 1 GPM IMERG satellite precipitation estimate improved as compared to TRMM TMPA-RT? *J. Geophys. Res. Atmos.* **2018**, *123*, 1694–1707. [CrossRef]

31. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]

32. Yuan, F.; Zhang, L.; Soe, K.M.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Applications of TRMM-and GPM-era multiple-satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in myanmar. *Remote Sens.* **2019**, *11*, 140. [CrossRef]

33. Wang, H.; Yong, B. Quasi-Global Evaluation of IMERG and GSMaP Precipitation Products over Land Using Gauge Observations. *Water* **2020**, *12*, 243. [CrossRef]

34. Anjum, M.N.; Ahmad, I.; Ding, Y.; Shangguan, D.; Zaman, M.; Ijaz, M.W.; Sarwar, K.; Han, H.; Yang, M. Assessment of IMERG-V06 Precipitation Product over Different Hydro-Climatic Regimes in the Tianshan Mountains, North-Western China. *Remote Sens.* **2019**, *11*, 2314. [CrossRef]

35. Fang, J.; Yang, W.; Luan, Y.; Du, J.; Lin, A.; Zhao, L. Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. *Atmos. Res.* **2019**, *223*, 24–38. [CrossRef]

36. Ouyang, L.; Yang, K.; Lu, H.; Chen, Y.; La, Z.; Zhou, X.; Wang, Y. Ground-based observations reveal unique valley precipitation patterns in the central Himalaya. *J. Geophys. Res. Atmos.* **2020**, *125*. [CrossRef]

37. Retalis, A.; Katsanos, D.; Tymvios, F.; Michaelides, S. Validation of the first years of GPM operation over Cyprus. *Remote Sens.* **2018**, *10*, 1520. [CrossRef]

38. Sungmin, O.; Foelsche, U.; Kirchengast, G.; Fuchsberger, J.; Tan, J.; Petersen, W.A. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6559–6572.

39. Cattani, E.; Merino, A.; Levizzani, V. Evaluation of Monthly Satellite-Derived Precipitation Products over East Africa. *J. Hydrometeorol.* **2016**, *17*, 2555–2573. [CrossRef]

40. Prakash, S.; Mitra, A.K.; Rajagopal, E.; Pai, D. Assessment of TRMM-based TMPA-3B42 and GSMaP precipitation products over India for the peak southwest monsoon season. *Int. J. Clim.* **2015**, *36*, 1614–1631. [CrossRef]

41. Yatagai, A.; Kawamoto, H. Quantitative estimation of orographic precipitation over the Himalayas by using TRMM/PR and a dense network of rain gauges. *Proc. Spie* **2008**, *7148*. [CrossRef]

42. Duncan, J.M.; Biggs, E.M. Assessing the accuracy and applied use of satellite-derived precipitation estimates over Nepal. *Appl. Geogr.* **2012**, *34*, 626–638. [CrossRef]

43. Yamamoto, M.K.; Ueno, K.; Nakamura, K. Comparison of Satellite Precipitation Products with Rain Gauge Data for the Khumb Region, Nepal Himalayas. *J. Meteorol. Soc. Jpn.* **2011**, *89*, 597–610. [CrossRef]

44. Derin, Y.; Anagnostou, E.; Berne, A.; Borga, M.; Boudevillain, B.; Buytaert, W.; Chang, C.-H.; Chen, H.; Delrieu, G.; Hsu, Y.C. Evaluation of GPM-era Global Satellite Precipitation Products over Multiple Complex Terrain Regions. *Remote Sens.* **2019**, *11*, 2936. [CrossRef]

45. Dinku, T.; Ceccato, P.; Grover-Kopec, E.; Lemma, M.; Connor, S.; Ropelewski, C. Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* **2007**, *28*, 1503–1526. [CrossRef]

46. Wang, S.; Liu, J.; Wang, J.; Qiao, X.; Zhang, J. Evaluation of GPM IMERG V05B and TRMM 3B42V7 Precipitation Products over High Mountainous Tributaries in Lhasa with Dense Rain Gauges. *Remote Sens.* **2019**, *11*, 2080. [CrossRef]

47. Gebregiorgis, A.S.; Hossain, F. Estimation of satellite rainfall error variance using readily available geophysical features. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 288–304. [CrossRef]

48. Khadka, N.; Zhang, G.; Thakuri, S. Glacial lakes in the Nepal Himalaya: Inventory and decadal dynamics (1977–2017). *Remote Sens.* **2018**, *10*, 1913. [CrossRef]

49. Nayava, J.L. Rainfall in Nepal. *Himal. Rev.* **1980**, *12*, 1–18.

50. Department of Hydrology and Meteorology (DHM); Government of Nepal. Observed Climate Trend Analysis of Nepal (1971–2014); Department of Hydrology and Meteorology: Kathmandu, Nepal. Available online: http://www.dhm.gov.np/climate/ (accessed on 1 October 2019).

51. Hamal, K.; Sharma, S.; Baniya, B.; Khadka, N.; Zhou, X. Inter-annual variability of Winter Precipitation over Nepal coupled with ocean-atmospheric patterns during 1987-2015. *Front. Earth Sci.* **2020**, *8*, 161. [CrossRef]

52. Shrestha, M.L. Interannual variation of summer monsoon rainfall over Nepal and its relation to Southern Oscillation Index. *Theor. Appl. Clim.* **2000**, *75*, 21–28. [CrossRef]

53. Talchabhadel, R.; Karki, R.; Parajuli, B. Intercomparison of precipitation measured between automatic and manual precipitation gauge in Nepal. *Measurement* **2017**, *106*, 264–273. [CrossRef]

54. Nie, Y.; Sheng, Y.; Liu, Q.; Liu, L.; Liu, S.; Zhang, Y.; Song, C. A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015. *Remote Sens. Environ.* **2017**, *189*, 1–13. [CrossRef]

55. Becker, A.; Schneider, U.; Meyer-Christoffer, A.; Ziese, M.; Finger, P.; Stender, P.; Heller, A.; Breidenbach, J. *GPCC Report for Years 2009 and 2010*; DWD: Offenbach, Germany, 2011.

56. Yatagai, A.; Arakawa, O.; Kamiguchi, K.; Kawamoto, H.; Nodzu, M.I.; Hamada, A. A 44-Year Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Sola* **2009**, *5*, 137–140. [CrossRef]

57. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. *NASA/GSFC Code* **2019**, *612*, 69.

58. Tan, J.; Huffman, G.J. Computing Morphing Vectors for Version 06 IMERG. Available online: https://gpm.nasa.gov/sites/default/files/document_files/MorphingInV06IMERG.pdf (accessed on 10 December 2019).

59. Legates, D.R.; Willmott, C.J. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Clim.* **1990**, *10*, 111–127. [CrossRef]

60. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Xie, P.; Yoo, S.-H. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theor. Basis Doc. (ATBD)* **2019**, *6*, 26.

61. Ushio, T.; Sasashige, K.; Kubota, T.; Shige, S.; Okamoto, K.; Aonashi, K.; Inoue, T.; Takahashi, N.; Iguchi, T.; Kachi, M.; et al. A Kalman Filter Approach to the Global Satellite Mapping of Precipitation (GSMaP) from Combined Passive Microwave and Infrared Radiometric Data. *J. Meteorol. Soc. Jpn.* **2009**, *87*, 137–151. [CrossRef]

62. Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Hirose, M.; Takayabu, Y.N.; Ushio, T.; Nakagawa, K. Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2259–2275. [CrossRef]

63. Shige, S.; Yamamoto, T.; Tsukiyama, T.; Kida, S.; Ashiwake, H.; Kubota, T.; Seto, S.; Aonashi, K.; Okamoto, K. The GSMaP Precipitation Retrieval Algorithm for Microwave Sounders-Part I: Over-Ocean Algorithm. *IEEE Trans. Geosci. Remote* **2009**, *47*, 3084–3097. [CrossRef]

64. Lu, D.; Yong, B. Evaluation and hydrological utility of the latest GPM IMERG V5 and GSMaP V7 precipitation products over the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 2022. [CrossRef]

65. Liechti, T.C.; Matos, J.P.; Boillat, J.L.; Schleiss, A.J. Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi River Basin. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 489–500. [CrossRef]

66. Thiemig, V.; Rojas, R.; Zambrano-Bigiarini, M.; Levizzani, V.; De Roo, A. Validation of Satellite-Based Precipitation Products over Sparsely Gauged African River Basins. *J. Hydrometeorol.* **2012**, *13*, 1760–1783. [CrossRef]

67. Bai, P.; Liu, X.M. Evaluation of Five Satellite-Based Precipitation Products in Two Gauge-Scarce Basins on the Tibetan Plateau. *Remote Sens. (Basel)* **2018**, *10*, 1316. [CrossRef]

68. Wang, W.; Lu, H.; Zhao, T.J.; Jiang, L.M.; Shi, J.C. Evaluation and Comparison of Daily Rainfall From Latest GPM and TRMM Products Over the Mekong River Basin. *IEEE J.-Stars* **2017**, *10*, 2540–2549. [CrossRef]

69. Feidas, H. Validation of satellite rainfall products over Greece. *Theor. Appl. Clim.* **2010**, *99*, 193–216. [CrossRef]

70. Li, Z.; Yang, D.W.; Hong, Y. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *J. Hydrol.* **2013**, *500*, 157–169. [CrossRef]

71. Wang, G.F.; Zhang, X.W.; Zhang, S.Q. Performance of Three Reanalysis Precipitation Datasets over the Qinling-Daba Mountains, Eastern Fringe of Tibetan Plateau, China. *Adv. Meteorol.* **2019**. [CrossRef]

72. Kansakar, S.R.; Hannah, D.M.; Gerrard, J.; Rees, G. Spatial pattern in the precipitation regime of Nepal. *Int. J. Clim.* **2004**, *24*, 1645–1659. [CrossRef]

73. Derin, Y.; Yilmaz, K.K. Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. *J. Hydrometeorol.* **2014**, *15*, 1498–1516. [CrossRef]

74. Bui, H.T.; Ishidaira, H.; Shaowei, N. Evaluation of the use of global satellite–gauge and satellite-only precipitation products in stream flow simulations. *Appl. Water Sci.* **2019**, *9*, 53. [CrossRef]

75. Yamamoto, M.K.; Shige, S.; Yu, C.-K.; Cheng, L.-W. Further improvement of the heavy orographic rainfall retrievals in the GSMaP algorithm for microwave radiometers. *J. Appl. Meteorol. Clim.* **2017**, *56*, 2607–2619. [CrossRef]

76. Ichiyanagi, K.; Yamanaka, M.D.; Muraji, Y.; Vaidya, B.K. Precipitation in Nepal between 1987 and 1996. *Int. J. Clim.* **2007**, *27*, 1753–1762. [CrossRef]

77. Lang, T.J.; Barros, A.P. An investigation of the onsets of the 1999 and 2000 monsoons in central Nepal. *Monthly Weather Rev.* **2002**, *130*, 1299–1316. [CrossRef]

78. Anjum, M.N.; Ding, Y.; Shangguan, D.; Tahir, A.A.; Iqbal, M.; Adnan, M. Comparison of two successive versions 6 and 7 of TMPA satellite precipitation products with rain gauge data over Swat Watershed, Hindukush Mountains, Pakistan. *Atmospheric Sci. Lett.* **2016**, *17*, 270–279. [CrossRef]

79. Grody, N.C.; Weng, F. Microwave emission and scattering from deserts: Theory compared with satellite measurements. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 361–375. [CrossRef]

80. Zeng, Q.; Wang, Y.; Chen, L.; Wang, Z.; Zhu, H.; Li, B. Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013. *Remote Sens.* **2018**, *10*, 168. [CrossRef]

81. Beck, H.E.; Van Dijk, A.I.; Levizzani, V.; Schellekens, J.; Gonzalez Miralles, D.; Martens, B.; De Roo, A. MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 589–615. [CrossRef]

82. Guo, H.; Chen, S.; Bao, A.; Hu, J.; Gebregiorgis, A.; Xue, X.; Zhang, X. Inter-comparison of high-resolution satellite precipitation products over Central Asia. *Remote Sens.* **2015**, *7*, 7181–7211. [CrossRef]

83. Yang, Y.; Tang, J.; Xiong, Z.; Dong, X. Evaluation of High-Resolution Gridded Precipitation Data in Arid and Semiarid Regions: Heihe River Basin, Northwest China. *J. Hydrometeorol.* **2017**, *18*, 3075–3101. [CrossRef]

84. Lakew, H.B.; Moges, S.A.; Asfaw, D.H. Hydrological evaluation of satellite and reanalysis precipitation products in the Upper Blue Nile Basin: A case study of Gilgel Abbay. *Hydrology* **2017**, *4*, 39. [CrossRef]

85. Dinku, T.; Hailemariam, K.; Maidment, R.; Tarnavsky, E.; Connor, S. Combined use of satellite estimates and rain gauge observations to generate high-quality historical rainfall time series over Ethiopia. *Int. J. Clim.* **2014**, *34*, 2489–2504. [CrossRef]

86. Derin, Y.; Anagnostou, E.; Berne, A.; Borga, M.; Boudevillain, B.; Buytaert, W.; Chang, C.H.; Delrieu, G.; Hong, Y.; Hsu, Y.C.; et al. Multiregional Satellite Precipitation Products Evaluation over Complex Terrain. *J. Hydrometeorol.* **2016**, *17*, 1817–1836. [CrossRef]

87. Satge, F.; Hussain, Y.; Bonnet, M.P.; Hussain, B.M.; Martinez-Carvajal, H.; Akhter, G.; Uagoda, R. Benefits of the Successive GPM Based Satellite Precipitation Estimates IMERG-V03,-V04,-V05 and GSMaP-V06,-V07 Over Diverse Geomorphic and Meteorological Regions of Pakistan. *Remote Sens. (Basel)* **2018**, *10*, 1373. [CrossRef]

Article

High Quality Zenith Tropospheric Delay Estimation Using a Low-Cost Dual-Frequency Receiver and Relative Antenna Calibration

Andreas Krietemeyer *, Hans van der Marel, Nick van de Giesen and Marie-Claire ten Veldhuis

Faculty of Civil Engineering, Delft University of Technology, 2628CN Delft, The Netherlands;
H.vanderMarel@tudelft.nl (H.v.d.M.); N.C.vandeGiesen@tudelft.nl (N.v.d.G.);
j.a.e.tenveldhuis@tudelft.nl (M.-C.t.V.)
* Correspondence: A.Krietemeyer@tudelft.nl

Abstract: The recent release of consumer-grade dual-frequency receivers sparked scientific interest into use of these cost-efficient devices for high precision positioning and tropospheric delay estimations. Previous analyses with low-cost single-frequency receivers showed promising results for the estimation of Zenith Tropospheric Delays (ZTDs). However, their application is limited by the need to account for the ionospheric delay. In this paper we investigate the potential of a low-cost dual-frequency receiver (U-blox ZED-F9P) in combination with a range of different quality antennas. We show that the receiver itself is very well capable of achieving high-quality ZTD estimations. The limiting factor is the quality of the receiving antenna. To improve the applicability of mass-market antennas, a relative antenna calibration is performed, and new absolute Antenna Exchange Format (ANTEX) entries are created using a geodetic antenna as base. The performance of ZTD estimation with the tested antennas is evaluated, with and without antenna Phase Center Variation (PCV) corrections, using Precise Point Positioning (PPP). Without applying PCVs for the low-cost antennas, the Root Mean Square Errors (RMSE) of the estimated ZTDs are between 15 mm and 24 mm. Using the newly generated PCVs, the RMSE is reduced significantly to about 4 mm, a level that is excellent for meteorological applications. The standard U-blox ANN-MB-00 patch antenna, with a circular ground plane, after correcting the phase pattern yields comparable results (0.47 mm bias and 4.02 mm RMSE) to those from geodetic quality antennas, providing an all-round low-cost solution. The relative antenna calibration method presented in this paper opens the way for wide-spread application of low-cost receiver and antennas.

Keywords: GNSS; GNSS antenna; receiver antenna calibration; relative calibration; Phase Center Variation; U-blox; goGPS; Zenith Tropospheric Delay; ZED-F9P

1. Introduction

The use of Global Navigation Satellite Systems (GNSS) data is well established for a range of professional and scientific applications, including atmospheric research. Precipitable Water Vapor (PWV) can be derived from radiosonde measurements, but also estimated with GNSS phase measurements processing techniques [1]. Zenith Tropospheric Delays (ZTDs) are estimated along with other parameters, such as the station position and receiver clock offset. The estimated ZTD is related to the refractive index of air, it can be used directly in atmospheric models, or ZTD can be used to estimate PWV. These values are of essential interest for accurate short-term weather forecasts.

The use of cost-efficient GNSS equipment is of great scientific interest to meteorology and atmospheric research. Traditionally, atmospheric research and meteorology made use of already existing GNSS infrastructures which were set-up for surveying and geodetic applications. Therefore,

low-cost GNSS equipment can be a good solution for regions where (commercial) high grade equipment is not readily available.

Analysing ZTDs or PWV from GNSS receivers has proven to be a valuable tool for analysing and forecasting extreme rainfall events [2]. Nowadays Numerical Weather Prediction (NWP) models exist that use these parameters in their data assimilation schemes to improve their rainfall predictions [3]. To estimate ZTD at existing GNSS receivers on the ground, numerous Analysis Centers (ACs) apply techniques such as Double-Differenced (DD) processing of large GNSS networks, or, the stand-alone Precise Point Positioning (PPP) technique [4]. In contrast to network or relative positioning strategies in general, PPP uses undifferenced GNSS observations and relies on precise satellite clock and positions to be available from the International GNSS Service (IGS). Dual-frequency observations are used to form the ionosphere-free linear combination, which removes the majority of the error associated with the delay in the ionosphere layer, but increases the noise by a factor of 3 [5]. In the past, estimating ZTD for meteorological purposes using cost-efficient equipment was only possible with single-frequency receivers, which required interpolation of the ionospheric delay from a network of dual-frequency receivers (e.g., References [6,7]). Recently, also low-cost dual-frequency receivers became available, which in combination with low-priced antennas, inspired the experiment in this paper. This experiment aims to investigate the potential use of low-cost dual-frequency receivers for ZTD estimation, something that was formerly only possible using expensive receiver equipment, or relied on existing geodetic networks to perform corrections. For this experiment different quality antennas, ranging from geodetic to mass-market, are deployed consecutively on a short-baseline. The new low-cost design of the receiver supplemented by a low-cost antenna may enable the cost-efficient estimation of high-quality tropospheric delays.

For high-precision applications, both satellite and receiver antenna corrections need to be applied [8]. These corrections are not only frequency dependent, but they also depend on the azimuth and elevation of the transmitting satellite. The delay caused by the antenna and the near field environment is not the same in every direction. Neither is it possible to find, for a specific frequency, a single point in the antenna for which the delay is the same for every direction. Therefore, a geometric center is something that does not really exists, and concepts like Antenna Phase Center (APC) are only approximate. These variations in antenna delay cause errors in the horizontal and vertical position, but also in the ZTD estimation. In order to correct the carrier phase signals for this effect, antenna calibrations are performed. The result of an antenna calibration is represented by a Phase Center Offset (PCO) with respect to a chosen Antenna Reference Point (ARP), and azimuth- and elevation dependent Phase Center Variations (PCV) with respect to the PCO. PCV and PCO corrections are provided per frequency. They are not independent, and should always be used together in high-precision positioning applications. Applications that require less precision can use only PCO values. On the other hand, for ZTD estimation, the PCO values are not important, and only the PCV values matter.

To obtain PCO and PCV corrections, relative and absolute antenna calibrations can be used. Absolute calibrations are typically expensive and involve sending the antenna to a calibration facility, while a relative calibration can be performed more easily by yourself. The relative calibration is performed over a short baseline, with on one side the antenna to be calibrated, and on the other side a reference antenna that has already been calibrated before (see e.g., Reference [9]). Double-differenced carrier phase residuals are used to compute the relative pattern between the antenna's. It uses the fact that over short baselines for DD observations most errors (e.g., satellite clock and atmospheric delays) are eliminated. The drawback of this calibration method is that it is a relative calibration. Since 2006 [10] it has been common for geodetic applications to use absolute antenna calibrations. Therefore, to obtain an absolute calibration for our antenna, we must add the absolute calibration for the reference antenna to the relative calibration. Absolute calibrations are typically obtained from measurements in an anechoic chamber (see e.g., Reference [11] or Reference [12]) or robotic field calibrations (see Reference [13]). Robotic calibrations have the advantage that they are not affected by multipath effects as in a relative correction.

While high quality geodetic antennas are typically manufactured with highest precision to enable the repetition of phase patterns on the receiving phase center, this is only partially the case for low-cost antennas. Companies specialized in antenna calibrations, such as Geo++ [14] in Germany, generate individual absolute antenna calibrations for geodetic antennas. Absolute antenna calibrations are typically supplied in an Antenna Exchange Format (ANTEX) file [15] that contains PCO as well as azimuth- and elevation dependent PCV for different frequencies and satellite systems. Various individual calibrations of the same antenna type are averaged to type mean calibrations and distributed, such as by the IGS. Using these antenna patterns is a standard practice in geodesy and essential for high precision positioning and also for tropospheric delay estimations. Several studies were performed that compare the differences between type mean and individual antenna phase center calibrations for example, by Araszkiewicz and Völksen (2016) [16], Schmid et al. 2005 [17] and Sidorov and Teferle 2016 [18]. They found a typical difference of only 2 mm horizontal and up to 4 mm vertical between type mean and individual antenna calibrations. For one antenna an offset of up to 17 mm in the vertical and 10 mm in the horizontal direction was observed. The vertical positioning performance gives an indication about the accuracy of the troposphere estimations since both parameters are correlated [19]. Only few studies were performed that analyze the impact of the antenna PCV on tropospheric parameter estimations. Ejigu et al. 2018 [20] investigated the impact of individual and type mean calibrations on Zenith Wet Delays (ZWDs) and tropospheric gradients [21]. They report a mean ZWD bias of 1.8 mm. A study by Pacione et al. 2017 [22] demonstrate similar results with ZTD errors in the mm range between utilizing individual and type mean calibrations.

The low-cost GNSS chips, being affordable and produced for the mass-market, are also of interest for scientific applications, but this only makes sense if also a low-cost antenna can be used. Geodetic antennas are generally characterized by special manufactured designs (e.g., choke ring) to suppress multipath induced effects. With low-cost antennas this is typically not the case. An analysis of low-cost antennas is especially important since they are expected to be particularly prone to multipath effects that are reflected in the antenna phase patterns.

In this paper we investigate the PPP ZTD performance of a recently introduced low-cost dual-frequency receiver connected to antennas of ranging quality (geodetic to mass-market) with and without applying relative antenna calibrations. Additionally, using L1-only data we investigate how well the (un-)corrected single-frequency data from the dual-frequency receiver can be used for meteorological applications. This is achieved by trimming the original data to L1-only datasets and generating a second frequency observation with the Satellite-specific Epoch-differenced Ionospheric Delay (SEID) model which is adjusted to the ionospheric delay based on a surrounding dual-frequency network. The paper is structured as follows—Section 2 describes the methodology, experimental setup and data analysis, Section 3 illustrates the PPP-derived ZTD results, Section 4 the discussion and Section 5 the conclusion.

2. Methods

Our experiment includes several steps to estimate the impact of different quality antennas on the ZTD performance using a low-cost dual-frequency receiver. The fundamental step upon which our investigation is based is a short-baseline analysis to perform a relative antenna calibration. Thereafter, different relative antenna calibration results are evaluated using the ZTD from two different PPP experiments. The experimental setup and overview of the data analysis procedure are described in Sections 2.1 and 2.2. The antenna calibration procedure is explained in Section 2.3. Section 2.4 describes the evaluation of the ZTD estimations with dual-frequency data in more detail. The antenna calibration impact on single-frequency data is covered in Section 2.5.

2.1. Experimental Setup

This subsection describes the test site, instrument setup, data handling as well as the receivers and antennas that were used in the experiment.

Our experiment consists of a series of consecutive short-baseline experiments using the International GNSS Service (IGS) station DLF1 as a base station and a U-blox ZED-F9P as rover. The DLF1 station is located on the rooftop of the Nederlands Metrology institute (NMi) in Delft. The station uses a Trimble NetR9 receiver, capable of tracking most GNSS signals, with a Leica AR25.R3 (LEIAR25.R3) chokering antenna with LEIT radome. The antenna has been calibrated by Geo++. The antenna is installed on the permanent marker with DOMES number 13502M009.

The rover receiver is an engineering sample of recently released low-cost dual-frequency receivers, that was obtained from the manufacturer for testing purposes. It covers the following frequencies: GPS L1C/A, L2C, GLONASS L1OF, L2OF, Galileo E1, E5b and BeiDou B1l, B2l and QZSS L1C/A, L2C. Notably missing GNSS signals are GPS L2P/Y, GPS L5 and Galileo E6. However, with the new generation GPS satellites, the L2C signal becomes rapidly available on almost all satellites and enables obtaining dual-frequency measurements from an increasing number of available GPS satellites. With up to 184 channels available, the receiver is capable of tracking two frequencies on each of the described satellite constellations and is still able to receive correction service data from augmented GNSS (e.g., Reference [23]).

The antennas of the rover are placed on a geodetic marker at a horizontal distance of approximately 10 m and 1.5 m height difference from DLF1. Different antenna types have been installed during consecutive periods from 15 February 2019 onwards. The investigated rover antennas cover a range of prices, starting with the lowest cost antennas U-blox ANN-MB-00 and Taoglas AQHA50 between about 50 and 100 euros, the middle price segment antennas Trimble AV28 and Trimble GA530 and antennas with a price of above 1000 euros, LEIAR25.R3 LEIT and TRM55971.00 NONE (also known as Trimble Zephyr 2 Geodetic). For these antennas, the mount point DOMES 13502M003 (GPS Mark 15) is used (Figure 1b). The antennas with hole-mount design (Trimble AV28) and without screw-hole at the bottom (U-blox ANN-MB-00) are mounted on top of a metallic rectangular extension bracket and also with a circular metallic ground plane with 10 cm diameter. For the other tested antennas, a tribrach with an adjustable circular level is used. The receiver itself is placed in a pelican case in the proximity and data logging is performed with a Raspberry Pi Zero on a local SD card (Figure 1a). Depending on weather conditions and observation time, antennas are switched after at least having recorded three full days of raw data. Data is transferred manually from the SD card for post-processing purposes.

The installation environment (regarding near-field effects and multipath) can be regarded as relatively clean. Both antenna positions (DLF1 and GPS Mark 15) are characterized by an unobstructed view over the full horizon. The time frame of antenna placements is depicted in Table 1. All antennas are active and, depending on the type, require different voltage as input. The antennas GA530, LEIAR25.R3 and TRM55971.00 require 12 V input voltage which cannot be supplied by the Raspberry Pi. Instead, the antennas are powered by a Septentrio receiver and the antenna signal is split to the U-blox ZED-F9P. The other antennas are working with voltage at or below 5 V that is supplied via the USB port of the Raspberry Pi and is considered as stable.
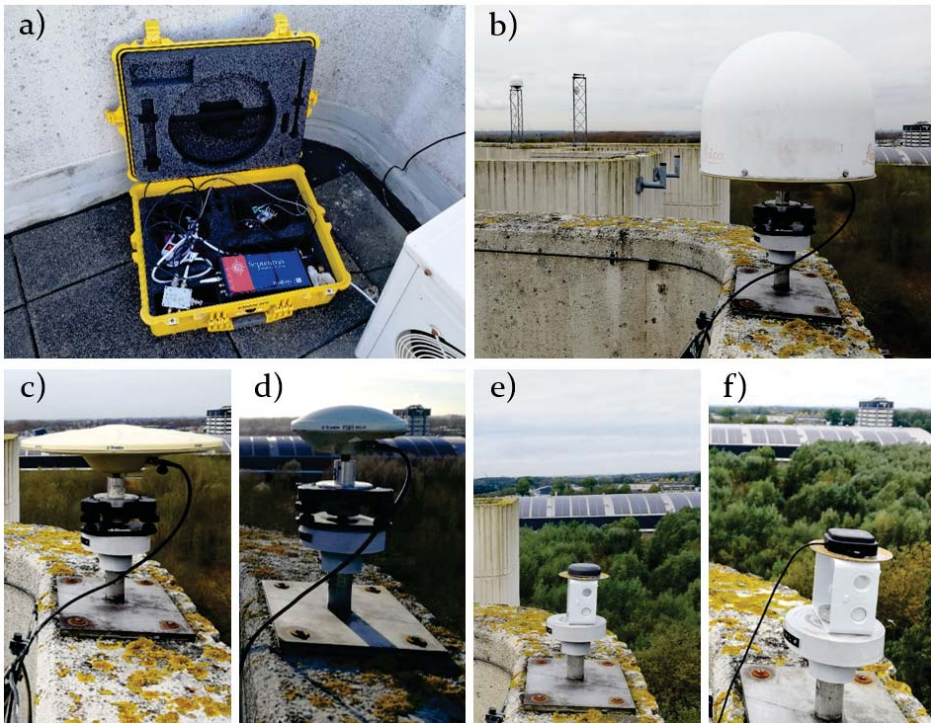
Figure 1. Waterproof pelican box rover equipment on the ground next to the marker (**a**). The pelican box contains the U-blox ZED-F9P, a Raspberry Pi Zero for data logging, power and a Septentrio receiver. The Septentrio receiver is used to provide power, through an antenna splitter, to the antennas: LEIAR25.R3 (**b**), Zephyr2 (**c**) and GA530 (**d**). No data collection or processing is performed with the Septentrio receiver. The verification setup is depicted in (**b**). It shows the baseline setup with two LEIAR25.R3 LEIT antennas. The antenna in the foreground is at the marker that is used for the investigation. In the background the radome and antenna of the base station DLF1 is visible. Both antennas are oriented North. The subfigures (**b**–**f**) illustrate the marker used for the investigation with different installed antennas. The antennas AV28 (**e**) and ANN-MB-00 (**f**) are depicted with a circular plane.

Table 1. Time frame and antenna descriptions of the antenna placements on GPS Mark 15 for the short-baseline experiments.

Start	End	DOYs	Antenna, Description
15 February 2019	20 February 2019	046–051	Trimble AV28 (stacked patch), rectangular bracket
20 February 2019	26 February 2019	051–057	Taoglas AQHA50 (stacked patch), rectangular bracket
26 February 2019	5 March 2019	057–064	U-blox ANN-MB-00 (stacked patch), rectangular bracket
5 March 2019	19 March 2019	064–078	Trimble GA530 (radome)
19 March 2019	26 March 2019	078–085	LEIAR25.R3 LEIT (choke ring)
26 March 2019	2 April 2019	085–092	Trimble TRM55971.00 NONE (resistive plane)
11 October 2019	15 October 2019	284–288	Trimble AV28 (stacked patch), circular plane
15 October 2019	18 October 2019	288–291	U-blox ANN-MB-00 (stacked patch), circular plane

The GA530 lost satellite tracking on L1 on 15 March 2019 17.23 UTC and one day later also L2 data was lost, presumably due to moisture in the antenna connector. Data from 15 March 2019 onwards is therefore discarded from the GA530 observations. The AQHA50 data was not processed due to

very low Signal-to-Noise (SNR) ratio, despite free-sky conditions, which we were unable to resolve. After quick and uncomplicated communication with the manufacturer the antenna could be returned and a replacement was provided. It was, however, not examined further in this experiment because of time limitations and practical considerations. A power outage in the Delft region, in the morning of Monday 25 March 2019, ended the data tracking for the LEIAR25.R3 antenna. Except for the AQHA50 antenna, sufficient data have been recorded for the remaining antennas for the analysis. Details about the data logging and conversion can be found in Appendix A.

The data is analysed in post-processing, with DLF1 as base station, and the antennas under investigation as rover. The rover data is available as 1-s daily RINEX 3 observation files. For the base station data, high-rate (1 Hz) RINEX3 data from the IGS station DLF1 were downloaded. High-rate 15-min RINEX 3 DLF1 observations were merged into daily files. Broadcast navigation data from the satellites were collected from the Crustal Dynamics Data Information System (CDDIS) [24]. For simplicity reasons and driven by the fact that most antenna calibrations are available for GPS, we used GPS-only data for our analysis.

2.2. Antenna Calibration and ZTD Evaluation Procedure

This subsection aims to provide a general overview of the antenna calibration and ZTD evaluation procedure. The experiment basically follows the steps illustrated in Figure 2. First, the antenna calibration is performed by retrieving residuals in a short baseline experiment, for which errors caused by the troposphere and ionosphere delays can be safely neglected. The original L1 + L2 RINEX data of base and rover were processed in RTKLIB, in static mode, to obtain the carrier phase residuals for each frequency, as well as azimuth (az) and elevation (el) angles of the corresponding satellites. The residuals, that can be considered as relative PCVs, together with IGS ANTEX type mean PCVs of the base station antenna (LEIAR25.R3 LEIT), were processed to create absolute PCVs which are saved in new ANTEX entries for each rover antenna. Details on the antenna types used in the experiment are provided in Table 1.
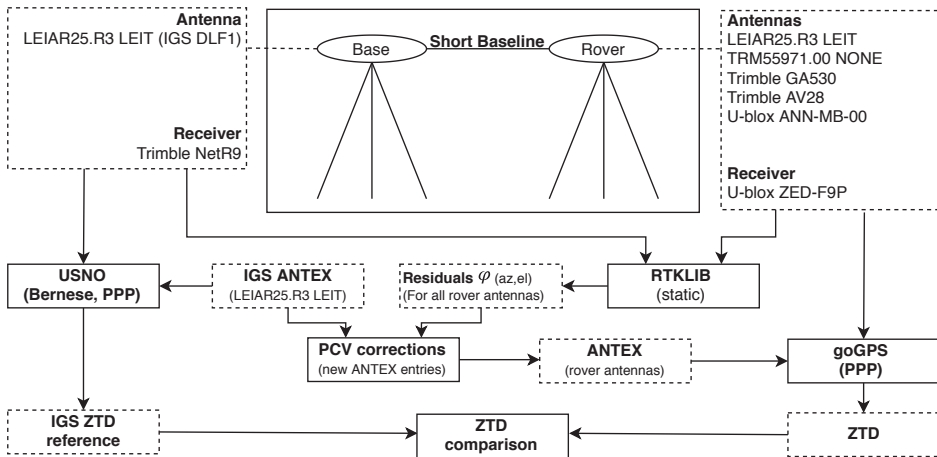
Figure 2. Steps in the Zenith Tropospheric Delay (ZTD) evaluation procedure. The top central box illustrates the short baseline with base station (DLF1) on the left and rover on the right. The base station DLF1 uses a Trimble NetR9 receiver and LEIAR25.R3 LEIT antenna, which did not change during the time frame of our experiment. The rover is illustrated on the right, with an U-blox ZED-F9P receiver and five antenna types applied consecutively.

Subsequently, the L1 + L2 RINEX data from the rover antennas, with the new ANTEX entries, were processed in goGPS (PPP-mode) to obtain absolute ZTD estimations. The obtained set of ZTD estimations for the low-cost receiver and different antennas was compared against the IGS final ZTD estimations by the US Naval Observatory (USNO) based on Bernese 5.2 in PPP mode from the IGS station DLF1. Further details on the antenna calibration and ZTD estimation procedures can be found in the following subsections.

2.3. Antenna Calibration and Retrieval of New ANTEX Entries

This subsection details the antenna calibration process. It is a fundamental step in our experiment and consists of an elevation-only and an azimuth and elevation dependent correction.

Near-field, multipath and antenna PCVs, summed up as multipath, affect all estimated parameters and residuals in the processing. Objects that are close to the antenna (near-field) affect the antenna patterns especially in the antenna PCO and azimuth- and elevation dependent PCVs. This effect cannot be covered in existing antenna calibrations without conducting the calibration procedure at the same location where the observation is taking place.

The main part of the antenna calibration consists of obtaining satellite phase residuals and is performed using RTKLIB. For each frequency, in our case GPS L1 and L2, the residuals are 'stacked' into elevation and azimuth bins. Two types approaches are used: (1) an elevation-only correction by averaging over all azimuth directions and (2) an azimuth and elevation dependent correction.

The GNSS processing engine RTKLIB [25] is open source and contains a range of Application Programs (APs) to perform real-time and post-processing precise positioning with GNSS data. It employs an Extended Kalman Filter (EKF) to obtain final differential solutions. In our analysis, the baseline processing was performed using the RTKLIB 2.4.3 Command-line User Interface (CLI) RNX2RTKP. Its functions are equivalent to the Graphical User Interface (GUI) program RTKPOST. In its standard configuration the program uses the highest elevation satellite as reference to obtain DDs. The residuals are written to a file for each satellite in view, with a zero value for the reference satellite, actually using a single difference (SD) format to store the DD residuals. For the analysis of the phase center variation we need however SD residuals. To obtain the SD residuals the average of the DD residuals (reference satellite included) must be subtracted from the DD residuals (again reference satellite included). This is the same as using the average DD residuals as reference. In order to use the average DD residuals as a reference, the source code was modified and recompiled. The observed measurement errors now had an expected zero mean of all satellite residuals on each frequency and not only for the highest elevation satellite. In our version the residuals were computed as in previous (legacy) RTKLIB releases. To avoid unintended behavior, the changes were discussed and changed in correspondence with the author of the package on github (see github issue: [26]).

Table 2 shows the processing settings used for the baseline analysis. Important processing options are the positioning solution to static (-p 3), elevation cut-off 5 degrees (-m 5), AR fix-and-hold (-h) and output residuals (-y 2). For Ambiguity Resolution (AR), the LAMBDA algorithm [27] is used within RTKLIB. The basic strategy in RTKLIB is to fix the ambiguities to integer after a float solution has been obtained. In our analysis we decided to use the fix-and-hold method. Further details on a conducted sensitivity case study can be found in Appendix B. The static processing option will strictly constrain the receiver movements for the observation period. It is preferred over a kinematic solution for the residual analysis since the rover antenna position does not change and the measurement errors should reflect this in the residuals and not in the estimated position. Driven by uncertainty about the implementation of applying ANTEX PCV corrections in RTKLIB, the antenna calibrations were not applied directly in the processing. Instead, their influence was analyzed separately. Each observation file of each antenna was processed and the generated output consisted of East, North and Up (ENU) components as well as frequency and azimuth and elevation dependent satellite residuals on code and phase measurements. For the relative antenna calibration, only the phase residuals were utilized.

Table 2. RTKLIB RNX2RTKP command line options

Option	Command-Line Parameter
Positioning solution: static	-p 3
Elevation cut-off: 5 degree	-m 5
AR: fix-and-hold	-h
AR validation threshold: 3	-v 3.0
Output: residuals	-y 2
Output: East, North, Up baseline	-a
Satellite systems: GPS-only	-sys G
Time format: YYYY/MM/DD hh:mm:ss	-t
Kalman filter: forward + backward	-c

To reduce noise and avoid outliers in the data, the residuals of each antenna are stacked over generally three selected DOYs (LEIAR25.R3: 79, 81, 83; TRM55971.00: 87, 88, 90; GA530: 65, 66, 70; AV28 with rectangular bracket: 47, 48, 49; AV28 with circular plane: 285, 286, 287; ANN-MB with rectangular bracket: 58, 61, 62) which were selected after a visible inspection. Only the ANN-MB-00 circular plane data is stacked over two days (DOYs 289 and 290) driven by the short observation period and evident outliers in the observed residuals.

In the first approach, the elevation-only phase patterns are visualised based on a first stacking, by taking the mean from all observations within 0.5 degree elevation bins. Afterwards, to correspond to the ANTEX standard, the stacked elevation-only phase residuals were averaged to 5 degree elevation bins using a moving-average filter. The resulting smoothed curve is shifted and start and end point (5 and 85 degrees) are substituted with the fitted data from a 10th order polynomial that was fitted over the 0.5 degree data. The resulting smoothed residual curve is the relative antenna calibration PCV (see also Section 3.1).

For the second approach, the azimuth and elevation dependent calibration, the number of residuals in each bin is not very large, and there will be many bins that are empty. To obtain meaningful azimuthal PCVs one should typically use longer observation time spans and rotate the antenna during the experiment in order to sample all azimuth directions. Since our experiment did not include rotating the antenna, and the observation period is only a few days, we do not expect very reliable azimuth corrections. To obtain azimuth dependent antenna patterns, the observed residuals were averaged over 5 degree elevation and azimuth bins and stacked over selected days. Since these measurements are generally noisy, the binned residuals are looped through the elevation bins and smoothed by fitting a 10th order polynomial to the data. To avoid evident outliers by fitting a polynomial to the data, gaps and missing data (especially in the North direction, see also Section 3.1) are filled by the nearest value from the current elevation bin. The 5 degree binned data used for the PCV calibration are taken from the resulting fitted polynomial curves.

For both approaches, to obtain absolute PCVs for each tested antenna, the estimated relative PCVs (averaged residuals) must be added to the absolute PCV of the base station antenna. For the absolute PCV of the base station antenna (LEIAR25.R3 LEIT) we used the IGS type mean azimuth and elevation dependent PCVs. Equation (1) shows how the rover antenna PCV is created:

$$\varphi_i^r(az, el) = \varphi_i^b(az, el) + \widetilde{\varphi}_i^{br}(az, el). \tag{1}$$

The superscripts *b* and *r* denote the base and rover, the subscript *i* the frequency, *az* and *el* the azimuth and elevation angles. By adding the estimated relative PCV of the rover ($\widetilde{\varphi}^{br}$) to the absolute base antenna PCV (φ^b), we obtained the absolute PCVs (φ^r) for the tested antenna. To use the newly generated PCVs in PPP tests, a new entry for each antenna was added to, or replaced in, the ANTEX file. Though individual calibrations are available for our base station antenna, we use the IGS I14.ATX ANTEX file that contains type mean calibrations for the base antenna PCV. A recent study

by Araszkiewicz et al., 2019 [28] investigated the height variation on the LEIAR25.R3 antenna with type mean and individual calibrations. They conclude that both methods show similar results.

For each approach (elevation-only and azimuth-elevation), a new ANTEX file is created with one new entry for each antenna that can be used for ZTD estimation using PPP. In this way, the antenna names given in the RINEX files do not have to be modified by instead supplying a respectively differently generated ANTEX file.

2.4. Evaluation of ZTD Estimates from Dual-Frequency PPP

This subsection describes the PPP processing configuration and ZTD evaluation procedure with dual-frequency GNSS measurements.

To estimate absolute ZTD values and to evaluate the newly obtained antenna PCVs, the open source tool goGPS [29] was used. Its ZTD estimations are comparable to existing reference datasets [6,7]. The ZTD results are computed for all available DOYs and compared to IGS final ZTD estimations from the IGS station DLF1 [24] in approximately 10 m distance. The closest available data point from the high rate ZTD estimations was used to match the IGS reference (5-min) interval.

We used goGPS PPP with the same configuration for all runs but with different ANTEX file configurations for the different antennas. The goGPS configuration used for our comparisons can be found in Table 3. The concept of the ZTD evaluation is summarized in Figure 3.

For all goGPS ZTD estimations, data from GPS-only satellites were used. An elevation cutoff angle of 10 degrees and ocean loading effects obtained from the FES2004 model [30] were applied. The IGS antenna calibration, final orbits, 30-s satellite clocks and earth rotation products were used [31].

The ZTD analysis consisted of up to five different cases. Two of our tested rover antennas, the LEIAR25.R3 LEIT and TRM55971.00 NONE, had already calibration entries in the IGS ANTEX file. This provided another set of PCVs for testing with PPP and serves as a verification of our estimations. The first case consists of the verification antennas (LEIAR25.R3 LEIT and TRM55971.00 NONE antennas). For this, a goGPS PPP run is performed using the original IGS I14.ATX ANTEX file. To investigate the impact of not using an ANTEX file at all, another run without the ANTEX file was performed in a second run. The third case applies only satellite PCO/PCV corrections by removing the receiver antenna-specific entries. The last two cases evaluate the elevation-only and azimuth-elevation corrected ANTEX entries obtained from the calibration step. One has to note that the full observation period (up to ten days observed data) is used as RINEX input for the ZTD estimation, while only two or three days are utilized to compute the PCVs.

Table 3. goGPS settings for Precise Point Positioning (PPP) ZTD estimation.

Type	Parameter
Observations	GPS-only
Elevation cutoff	10°
Ocean loading	FES2004
Observation weighting	same weight for all observations
Code observation error threshold	30 m
Phase observation error threshold	0.05 m
Code least-squares estimation error st. dev. threshold	40 m
Clock & orbits	IGS Final
Troposphere modeling	Saastamoinen (with GPT model)
Troposphere mapping function	GMF
Sampling interval	30-s
Antenna calibration	IGS, own, none
Kalman filter reset	no (seamless)
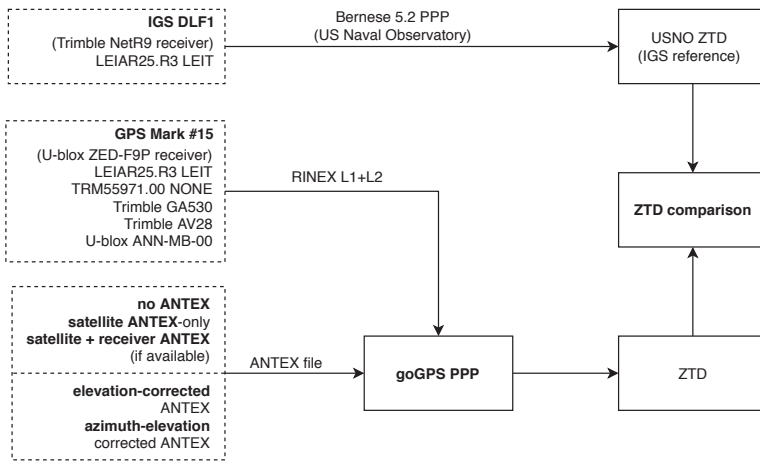
Figure 3. PPP ZTD evaluation concept. Original dual-frequency RINEX files are utilized as file input. Several runs with different antennas and ANTEX configurations are tested. The absolute ZTD estimations are evaluated against International GNSS Service (IGS) final ZTD estimations from DLF1.

If one runs PPP software, for example, goGPS, over which one has full control, one can use the newly generated ANTEX entries directly in the software. However, when PPP is performed in a server environment, one does not have the possibility to modify the ANTEX files. In these cases it is more convenient to correct the original RINEX phase data for the antenna patterns. To achieve this, we corrected the original RINEX phase data for the PCV estimations retrieved from the calibration step by the following formula:

$$\tilde{L}_i = L_i - \frac{\varphi_i^r(el, az)}{\lambda_i}, \tag{2}$$

where $\tilde{L}_i$ is the corrected phase data on frequency i and L_i the original RINEX phase observation. $\varphi_i^r(el, az)$ is the azimuth and elevation dependent antenna phase pattern on frequency i. λ_i denotes the wavelength on each frequency. A comparison between ANTEX-corrected and RINEX-corrected PPP ZTD comparisons resulted in no or negligible differences.

2.5. Evaluation of ZTD Estimates from Single-Frequency PPP

This subsection illustrates the single-frequency ZTD evaluation using the Satellite-specific Epoch-differenced Ionospheric Delay model (SEID) from a surrounding dual-frequency network. PPP uses the ionosphere-free linear combination to eliminate the first order effect of the ionospheric delay error. The drawback of this method is that the errors and multipath on L1 is increased by a factor 2.546, and on L2 by a factor 1.546. The effect is that, when the errors are not correlated, the noise of the ionosphere free linear combination is increased by a factor 3. Also for antennas that are sensitive to multipath, or are placed in a multipath-prone environment, this effect may cause the performance to decrease significantly. In these cases the performance of using L1-only measurements may outperform the dual-frequency ionosphere-free linear combination results (see also Reference [32]).

Instead, the SEID algorithm [33], combined with L1 data, generates a synthesized L2 measurement by utilizing data from a network of existing dual-frequency receivers. The synthesized signal contains in principle the same information as L1 but is adjusted to the ionospheric delay on that frequency based on the dual-frequency measurements from the surrounding network. This effectively reduces the noise effect. As part of the evaluation, we apply the SEID model on L1-data using a network of three receivers. By removing the second frequency observations from the RINEX data and correcting the L1 data for elevation and azimuth-elevation patterns, we also investigate the performance of the

SEID algorithm on L1-corrected data. Though stations in the close proximity to DLF1 are available, the outside station network distances are chosen between 55 and 130 km to simulate a more realistic scenario to densify existing networks. The SEID evaluation process is depicted in Figure 4.
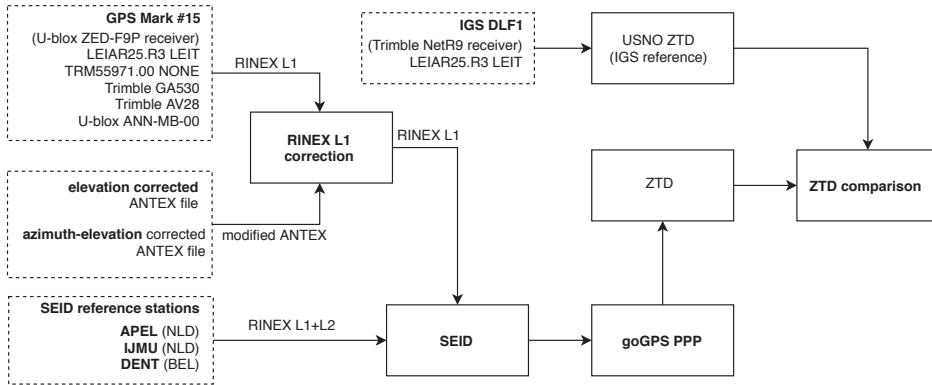
Figure 4. Satellite-specific Epoch-differenced Ionospheric Delay model (SEID) PPP methods. The original (dual-frequency) RINEX data is trimmed to L1-only observations and the RINEX correction after Equation (2) is utilized. SEID is applied with the stations APEL, IJMU and DENT. The subsequent PPP process is conducted once with the original I14.ATX file and once without the LEIAR25 and TRM55971.00 entries, hence providing satellite Phase Center Offset (PCO)/Phase Center Variation (PCV) corrections only. The resulting absolute ZTD estimations are evaluated against final IGS ZTD estimations from DLF1.

The following section describes the results of the experiment by analyzing the satellite phase residuals and ZTD estimations.

3. Results

3.1. Residuals Analysis and PCV Estimation to Obtain Corrected ANTEX Entries

First, we analyzed the residuals as a crucial step to correct for existing phase patterns caused by the receiving antenna. The residuals are derived from the static RTKLIB solution. The RTKLIB output contains the residuals for each satellite, epoch and frequency on code and phase. Table 4 shows the mean residuals and their respective RMSE.

The LEIAR25.R3 LEIT antenna demonstrated the lowest phase RMSE on both frequencies. It is also evident that the L2 phase residuals are generally higher than the L1 residuals for all antennas. The Trimble AV28 and U-blox ANN-MB-00 show generally higher phase RMSE values on both frequencies compared to the other antennas. The phase RMSE is considerably smaller when using a circular metallic ground plane. The antennas in Table 4 are ordered by approximated acquisition costs. A quality difference with generally lower phase residuals for the upper price category antennas compared to the less expensive ones is evident. As known from GNSS basics, the code residuals have higher deviations than the phase measurements. The mean code error is not zero because they are less precise and the mean single-difference phase residuals are utilized as reference in computing the double-differences. Though interesting, the code residuals are less important for the analysis because precise applications employ phase measurements. L1 and L2 mean phase residuals averaged over the elevation angle are shown in Figure 5.

Table 4. Mean and Root Mean Square Errors (RMSE) of code and phase residuals from the static solutions of the short baseline experiments. The phase mean is zero for all antennas since the mean of the single-differenced residuals are used as reference for the DD analysis in the modified RTKLIB application.

Antenna	Code RMSE (mm)		Code mean error (mm)		Phase RMSE (mm)	
	L1	L2	L1	L2	L1	L2
LEIAR25.R3 LEIT	508.36	522.46	−0.02	0.00	3.61	4.14
TRM55971.00 NONE	513.57	551.96	0.01	0.14	4.22	5.25
Trimble GA530	540.35	541.41	0.00	−0.02	3.99	5.42
Trimble AV28 (rectangular bracket)	644.51	578.63	0.08	0.11	7.70	8.93
Trimble AV28 (circular plane)	473.30	494.89	0.00	0.00	3.94	5.70
U-blox ANN-MB-00 (rectangular bracket)	632.60	599.06	-0.02	0.00	7.00	9.13
U-blox ANN-MB-00 (circular plane)	515.49	492.47	0.00	0.07	4.28	6.32
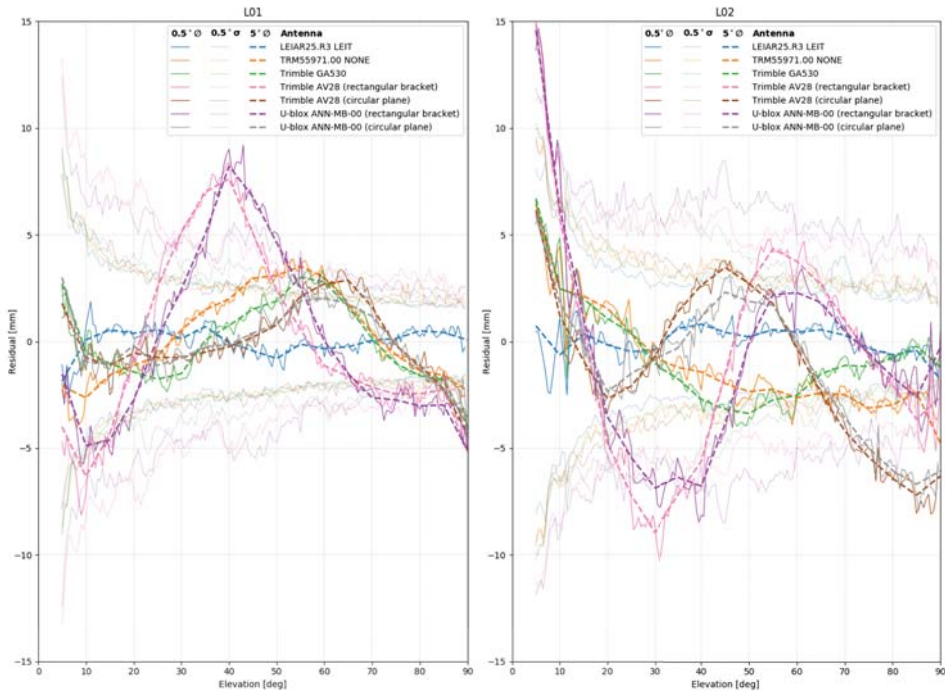
Figure 5. Elevation based antenna pattern on the phase residuals. L1 residuals are depicted left and L2 on the right. The continuous lines (0.5°∅) show the mean residual over 0.5 degree elevation bins, the dotted lines (0.5°σ) in the background its respective standard deviation and the dashed line (5°∅) the smoothed 5 degree bins corresponding to the ANTEX format.

Figure 5 shows the averaged elevation-based pattern for L1 and L2 observations with all tested antennas using a bin size of 0.5 degrees (continuous line), its standard deviations (thin dotted line in the background) and the residuals smoothed over 5 degree bins to correspond to the ANTEX format (thick dashed line). Generally apparent are higher variations at lower elevation angles for all antennas and a higher standard deviation of L2 compared to L1. A clear signal is evident for the Trimble AV28 and U-blox ANN-MB-00 antennas using the rectangular metal bracket as base. This effect can be seen on L1 with satellites between 30 and 50 degrees and on L2 at about 25 and 40 degrees. Since these results evoked further investigation, the experiment was repeated with a circular ground plane. The identified pattern could be reduced when using a circular ground plane. However, on L2 the observed variations above 60 degrees are noticeable greater compared to the other antennas. Evident is that both antennas, although being produced by a different manufacturer and having an inherently different design, demonstrate a comparable elevation-based pattern. The reference antenna LEIAR25.R3 LEIT demonstrated no clear elevation-based phase bias at any elevation angle since both, base and rover antenna are of the same model and the resulting residual error in differential analysis mostly cancels out. The antennas GA530 and TRM55971.00 also demonstrate a slightly visible mean phase pattern.

Typically, azimuth dependent residuals are stacked over several weeks or months of data or over several days by rotating the antenna by some degree. This was not performed in this study. Instead, with the limited observation days available we stacked the azimuth and elevation dependent residuals over the selected observation days. Since the antenna phase residuals of AV28 and ANN-MB-00 with different mountings evokes additional attention, Figure 6 compares the elevation and azimuth dependent L1 and L2 phase residuals for the reference antenna LEIAR25.R3 LEIT and the antennas Trimble AV28, U-blox ANN-MB-00 with rectangular bracket and circular plane.

Figure 6 demonstrates the performance of a reference antenna (LEIAR25.R3 LEIT) compared to the two low-cost antennas Trimble AV28 and U-blox ANN-MB-00 with a rectangular bracket and a circular plane as base. Note that the zenith angle is used instead of elevation angle to correspond with the ANTEX format and more intuitive interpretation of the polar plots. Generally noteworthy are the fewer satellite tracks on L2 which is caused by not all GPS satellites yet transmitting the L2C signal. As previously illustrated in the elevation based residuals, the LEIAR25.R3 antenna (Figure 6a) depicts generally low residuals over the full horizon with only a few higher values close to the ground. The patterns from the antennas AV28 and ANN-MB-00 are characterized by negligible differences between them. Noticeable differences are slightly lower residuals at low elevation in South-West (225°) direction on L1 for the AV28 (Figure 6b left) compared to the ANN-MB-00 (Figure 6d left) antenna using a rectangular bracket. Evident are also the strong negative residuals on L1 at low elevation in North (0°) direction and the stronger positive residuals on L2 in the same direction. The described L1 pattern at about 45 degree elevation (AV28 rectangular and ANN-MB-00 rectangular in Figure 5) is also visible with the strongest signal in East and West direction. The errors using a circular plane are generally smoother for both antennas. Regions with increased residuals are, however, still evident. For further processing, the elevation and azimuth dependent residuals were averaged in 5 degree bins to comply with the ANTEX standard. To overcome possible outliers in the data and since only limited observation days were used, we also applied a polynomial fit over the azimuth dependent signal. Based on the base antenna's (LEIAR25.R3 LEIT) ANTEX PCV entries, the elevation and azimuth-elevation dependent phase residuals were used to generate absolute PCVs to append or replace existing ANTEX entries for the tested antennas.
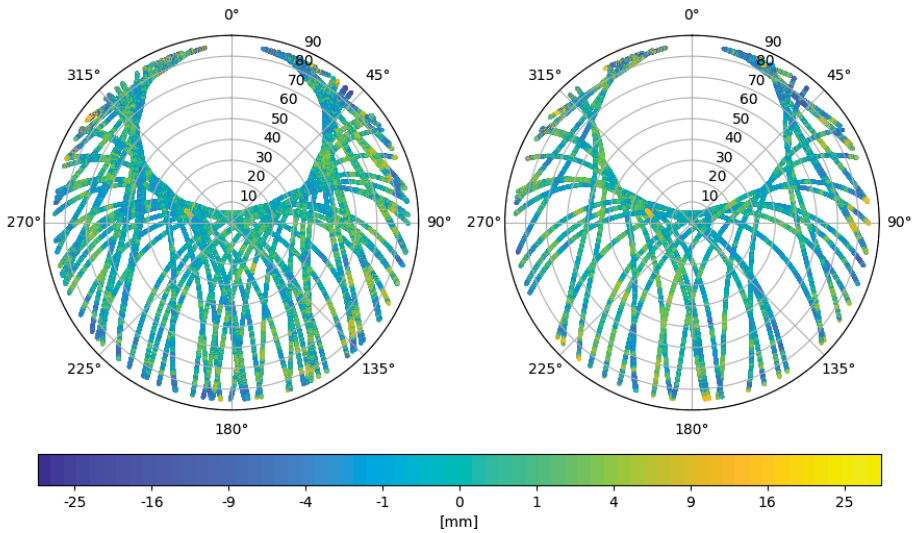
(**a**) Satellite phase residuals on L1 (top left) and L2 (top right) for the reference antenna LEIAR25.R3 LEIT.
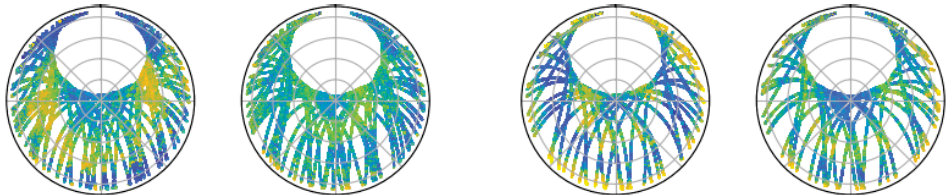
(**b**) AV28 L1 with a rectangular bracket (left) and a circular plane (right)

(**c**) AV28 L2 with a rectangular bracket (left) and a circular plane (right)
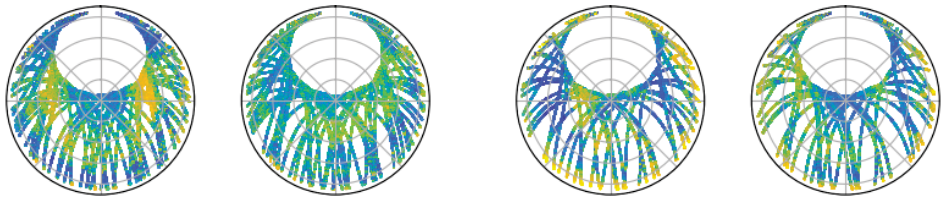
(**d**) ANN-MB-00 L1 with a rectangular bracket (left) and a circular plane (right)

(**e**) ANN-MB-00 L2 with a rectangular bracket (left) and a circular plane (right)

Figure 6. Satellite phase residuals on L1 (left) and L2 (right) for the selected antennas (**a**) LEIAR25.R3 LEIT (top frame), (**b**) AV28 L1, (**c**) AV28 L2, (**d**) ANN-MB-00 L1 and (**e**) ANN-MB-00 L2. The figures (**b**–**e**) depict on the left the phase residuals using a rectangular metallic bracket as base and on the right with a circular metallic plane of 10 cm diameter.

3.2. Evaluation of ZTD Estimates from Low-Cost Dual-Frequency Receiver and Different Quality Antennas

Table 5 shows the ZTD bias and RMSE values, computed without supplying an ANTEX file (no ANTEX), applying only satellite PCO and PCV corrections (Only satellite PCO/PCV), using the original IGS ANTEX entries (IGS ANTEX), elevation-only corrected ANTEX entries (El. Corrected) and azimuth-elevation dependent corrected ANTEX entries (Az.-el. Corrected). For the latter two runs, the existing ANTEX entries from the original IGS ANTEX file were removed and replaced with the generated antenna corrections from our experiment. For all evaluations, the IGS final ZTD estimations from the IGS station DLF1 located at approximately 10 m distance served as reference.

Table 5. ZTD estimation based on antenna PCV corrections with no ANTEX corrections, only satellite PCO/PCV corrections, original IGS type mean ANTEX, elevation corrected ANTEX and azimuth-elevation corrected ANTEX entries on the reference antennas LEIAR25.R3 LEIT and TRM55971.00 NONE. Bias and RMSE refer to the PPP ZTD estimations compared to the IGS ZTD reference.

Antenna	(mm)	No ANTEX	Only Satellite PCO/PCV	IGS ANTEX	El. Corrected	Az.-el. Corrected
LEIAR25.R3 LEIT	Bias	7.50	5.96	1.15	1.08	1.59
	RMSE	15.86	8.02	4.77	4.67	4.64
TRM55971.00 NONE	Bias	−0.89	−1.30	−0.98	−1.17	−2.26
	RMSE	14.20	3.83	3.58	3.59	3.85

The PPP runs without supplying an ANTEX file yielded high RMSE of about 15 mm for both antennas and a bias of 7.5 mm for the LEIAR25.R3 LEIT antenna. No significant bias is evident for the TRM55971.00 NONE antenna. Applying only the satellite PCO and PCV corrections, the bias of the LEIAR25.R3 antenna slightly decreased to about 6 mm and the RMSE to about 8 mm. The bias of the TRM55971.00 antenna remained at about the same level while the RMSE decreased drastically to about 4 mm which already agrees well with the expected standard deviation of about 4 mm from the official IGS ZTD final product [3,34]. Both reference antennas performed well using the original IGS ANTEX file resulting in no significant biases. The RMSE is in an expected range of about 4 mm. Replacing the existing ANTEX entries with the generated ANTEX entries resulted in a similar performance to the IGS reference results with RMSE values between 3 and 5 mm. Remarkably, compared to the LEIAR25.R3 antenna, the bias of the TRM55971.00 antenna increased up to about −2.3 mm in the azimuth- elevation corrected results.

The same evaluation is done for the antennas without official ANTEX entries, yet without the IGS receiver antenna corrections that are unavailable for this case. Table 6 shows the results for the ANTEX corrected PPP-based ZTD estimations with the antennas GA530, AV28 and ANN-MB-00.

Table 6. ZTD estimation based on antenna corrections with no ANTEX corrections, only satellite PCO/PCV corrections, elevation corrected ANTEX and azimuth-elevation corrected ANTEX entries on the antennas Trimble GA530, Trimble AV28 and U-blox ANN-MB-00. Bias and RMSE refer to the PPP ZTD estimations compared to the IGS ZTD reference.

Antenna	(mm)	No ANTEX	Only Satellite PCO/PCV	El-Corrected	Az-el-Corrected
Trimble GA530	Bias	3.65	−0.93	−2.21	−3.65
	RMSE	13.79	5.05	4.31	5.06
Trimble AV 28	Bias	−19.35	−24.04	−3.91	−2.86
(rectangular bracket)	RMSE	23.59	24.82	6.13	5.10
Trimble AV 28	Bias	12.14	9.54	2.61	2.83
(circular plane)	RMSE	18.02	10.86	5.01	4.79
U-blox ANN-MB-00	Bias	−14.54	−19.55	−3.34	−3.58
(rectangular bracket)	RMSE	20.19	20.59	6.64	5.44
U-blox ANN-MB-00	Bias	6.47	5.26	−0.47	−0.52
(circular plane)	RMSE	15.32	7.23	4.02	3.77

By supplying no ANTEX file or only correcting for the satellite antenna PCO and PCV, the results in Table 6 depict a generally better ZTD quality for the Trimble GA530 antenna compared to the Trimble AV28 and U-blox ANN-MB-00 antennas. A considerable difference is evident when using a circular ground plane and a rectangular bracket for the latter two. While the bias and RMSE of the data with rectangular brackets remained at a constant high level (>20 mm), employing a circular plane reduced the errors down to about 7 and 10 mm RMSE. Remarkably, applying the elevation-only or azimuth-elevation dependent ANTEX entries for these antennas resulted in a significant improvement for both antennas and mounting types. The offsets are lowered to between −0.47 and −3.91 mm and

the ZTD RMSE are between 3.77 and 6.64 mm. Whilst the RMSE of the Trimble GA530 remained at the same level of about 5 mm, the bias increased slightly compared to the satellite-only PCO and PCV corrections.

3.3. ZTD Evaluation Using Single-Frequency PPP with SEID Modelling

Table 7 shows ZTD results using L1 data combined with SEID modelling of L2. The Table shows results using the original RINEX L1 data, elevation-corrected L1 data and azimuth-elevation corrected L1 data as input. The IGS ANTEX file is used to provide the satellite PCO/PCVs. In case of the original RINEX L1 data, the receiver PCO/PCV is used as well in case there is an entry in the IGS ANTEX file. In case of the elevation and/or azimuth corrected data the RINEX L1 data is modified for the L1 PCV (no receiver entries from ANTEX are used).

Table 7. ZTD PPP results using original L1 data, elevation-corrected L1 data and azimuth-elevation corrected L1 data together with the original IGS ANTEX type mean entries using SEID. For the reference target antennas only satellite PCO/PCV ANTEX corrections were utilized. For the SEID processing, the Dutch Permanent GNSS Array (DPGA) stations APEL (Apeldoorn, Netherlands) and IJMU (IJmuiden, Netherlands) as well as the EUREF station DENT (Dentergem, Belgium) were used. The stations are located approximately 110, 55 and 130 km from the experimental setup.

Antenna	(mm)	Original L1 and IGS ANTEX	El-Corrected L1 and IGS ANTEX	Az-el-Corrected L1 and IGS ANTEX
LEIAR25.R3 LEIT	Bias	2.20	2.22	2.67
	RMSE	5.14	5.10	5.31
TRM55971.00 NONE	Bias	1.72	1.69	1.69
	RMSE	4.68	4.66	4.66
Trimble GA530	Bias	1.41	1.55	0.94
	RMSE	4.08	4.13	3.91
Trimble AV 28	Bias	−5.08	2.26	0.75
(rectangular bracket)	RMSE	7.58	5.96	5.42
Trimble AV 28	Bias	3.38	3.85	2.89
(circular plane)	RMSE	5.27	5.56	4.90
U-blox ANN-MB-00	Bias	−6.02	−0.36	−1.84
(rectangular bracket)	RMSE	7.34	4.16	4.44
U-blox ANN-MB-00	Bias	−0.40	0.22	−0.78
(circular plane)	RMSE	4.23	4.19	4.28

Applying SEID and using the original RINEX L1 data and IGS ANTEX file, the overall performance of the geodetic quality antennas (LEIAR25.R3 LEIT and TRM55971.00 NONE) remained at a high level. It is also evident that noise and bias were significantly decreased for the antennas Trimble AV28 and U-blox ANN-MB-00 for both mounting types compared to the satellite-only PCO and PCV corrected results from Table 6. The bias was almost zero for the U-blox ANN-MB-00 (circular plane) and −6 mm for the rectangular bracket dataset. The AV28 data were slightly more biased with values of about 3 mm for the circular plane data and −5 mm using the rectangular bracket. The RMSE values were between about 4 mm and 7.5 mm. Notably, the ANN-MB-00 (circular plane) dataset is already comparable to the reference results from Table 6. Applying elevation and azimuth-elevation dependent corrections to the RINEX L1 observations further improved the results with biases between almost zero and 4 mm and RMSE values of about 4 to 6 mm ZTD. The performance of the higher grade antennas (LEIAR25.R3 LEIT, TRM55971.00 NONE and Trimble GA530) remained at a stable level which is similar to the results from Table 6.

The results from all conducted evaluations are summarized in the boxplots shown in Figure 7. The boxplots show the differences between the IGS ZTD reference data and the obtained goGPS PPP ZTD estimations. The boxplots on the left depict the results using the original L1 + L2 RINEX data with the original IGS ANTEX, without supplying ANTEX corrections and using only satellite PCO and PCV. One should note that these graphs in the left box illustrate the ZTD results with the original

recorded dual-frequency data and different IGS ANTEX constellations. It is evident that supplying the original IGS ANTEX corrections resulted in the smallest errors, but these are not available for the low cost antennas. It also shows that applying no ANTEX corrections to the data yield high standard deviations and offsets in the ZTD estimations. Supplying only satellite PCO/PCV corrections provide reasonable results for the antennas TRM55971.00 NONE and GA530. The antennas LEIAR25.R3 LEIT, AV28 (circular) and ANN-MB-00 (circular) are at a comparable level. Although the standard deviation decreased for the data from the ANN-MB-00 and AV28 antennas, a significant bias was still present for the rectangular ground plate datasets.
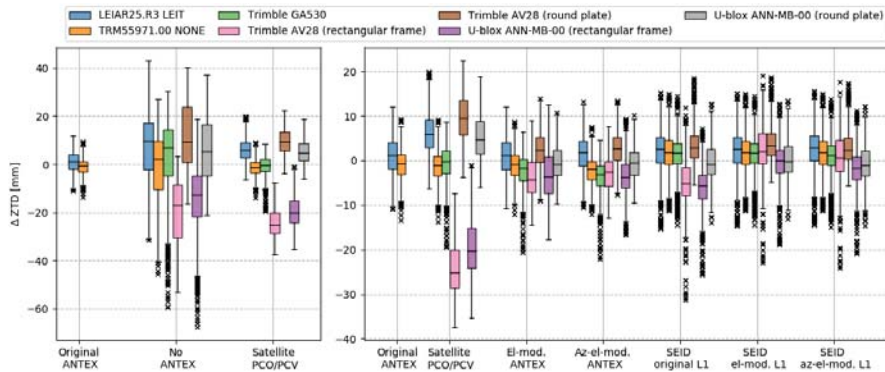
Figure 7. Boxplots of goGPS PPP ZTD estimations against IGS reference data from DLF1. The figure on the left depicts the combinations with the original dual-frequency RINEX data using the original IGS ANTEX file, supplying no ANTEX file at all, or only satellite PCO/PCV corrections. The boxplots on the right illustrate the references (original ANTEX, only satellite PCO/PCV), the results with original RINEX L1 + L2 and modified ANTEX entries (elevation modified ANTEX, azimuth-elevation modified ANTEX) and SEID results with original, elevation corrected and azimuth-elevation corrected L1 data.

The boxplots on the right in Figure 7 depict the reference data (original ANTEX and satellite PCO/PCV) and summarize the results for the corrected datasets (the first four groups) and applied SEID experiments (the rightmost three groups). One can note that the corrected data illustrate values in the range between −30 and 20 mm compared to −70 and 45 mm when not using the IGS ANTEX file. The presented results demonstrate a similar performance of all antennas with their most significant differences being the biases. Notably, the biases of the rectangular and circular mounting types are of opposite sign using only satellite PCO/PCV corrections. After correction, the biases are significantly reduced for the LEIAR25.R3 LEIT antenna as well as for the low-cost antennas AV28 and ANN-MB-00 using both, rectangular bracket and circular plane. It also shows that results for the single-frequency SEID experiments are comparable to the dual-frequency results.

4. Discussion

4.1. Residual Analysis

The presented residual analysis provides elevation- and azimuth-elevation phase patterns on both frequencies for each rover antenna. The analysis proves that antenna dependent residual phase patterns are present, and certain anomalies, for example, multipath signals presumably caused by the rectangular plate (AV28 and ANN-MB-00), can be exposed. The phase patterns can be modelled to provide corrections and/or a relative antenna calibration for our further investigations. The relative calibrations can be turned into absolute antenna calibrations in the international ANTEX format by averaging over 5 degree elevation and azimuth bins, and adding the relative calibration to the absolute calibration of the reference antenna. An important point of discussion is whether the full

azimuth-elevation pattern should be used, or the elevation only pattern. The elevation only pattern is more robust and less sensitive to possible outliers than the azimuth-elevation pattern, it is also computationally less expensive to compute. Three days of data is enough to compute the elevation only pattern. To compute the azimuth-elevation pattern we advise to use more days, and/or rotate the antenna, to fill the gaps in the azimuth patterns. Additional satellite tracks over the horizon can also be obtained by utilizing multiple satellite constellations sharing the same center frequency (e.g., Galileo E1 and GPS L1). However, it introduces additional unknowns (i.e., estimation of an additional clock term) and, in the latitude of the executed experiment, still leaves the gap in the North. Consequently, a carefully conducted rotation of the antenna and accounting for it during the processing can be used to fill gaps in the azimuth pattern and at the same time average potential multipath effects. The result we obtained for the elevation only patterns were comparable to our limited azimuth-elevation patters, thus, for three days calibration time, we recommend to use the elevation only patterns. The procedure for azimuth-elevation patterns needs some further investigation.

4.2. Absolute ZTD Estimations

In order to evaluate the performance of our antenna calibrations on the absolute ZTD estimation several PPP computations were performed. The rover antennas were divided in two groups. The first group is formed by the LEIAR25.R3 LEIT and TRM55971.00 NONE from which we know the antenna patterns from existing type mean IGS ANTEX entries, and which provide an excellent reference for our own calibrations, as well as scenarios in which no or satellite-only antenna calibrations would be available. The results are summarized in Table 5 and Figure 7. Using no ANTEX file at all, the ZTD RMSE compared to the IGS reference is about 15 mm for both antennas in contrast to about 4 mm with applying the IGS PCO and PCV corrections. Supplying only satellite antenna corrections decreases the ZTD RMSE to 8 mm (LEIAR25.R3 LEIT) and to about 4 mm (TRM55971.00 NONE). It also shows that the LEIAR25.R3 LEIT antenna, caused by the antenna design and radome, has a phase pattern that deviates significantly from a sphere, resulting in a bias of about 6 mm when applying only satellite antenna corrections (and ignoring the receiver corrections). This makes it particularly important to correct for this error with this antenna type. The TRM55971.00 NONE antenna pattern suggests to have only little impact on the ZTD estimation with almost identical results comparing the original IGS ANTEX data results. It suggests that the PCV of this antenna has a spherical pattern. Applying the antenna PCV corrections from the conducted short baseline analysis demonstrated comparable results for the elevation-based ANTEX correction and slightly less precise results for the azimuth-elevation corrections. Especially the bias increased which suggests that more outliers are present in the azimuth analysis which is presumably caused by the lack of observing days to perform an azimuthal correction with the utilized calibration method. Except for the minor bias, the ZTD RMSE with self-calibrated antenna pattern entries are comparable to the results using the IGS ANTEX calibrations, and comparable to the IGS reference ZTD estimations.

The second group of antennas consisted of the GA530, AV28 and ANN-MB-00. Since no official antenna calibrations are available for these antennas, the performance could not be evaluated directly against official calibrations. However, the IGS supplies official ZTD estimations from DLF1 to which the results can be compared. Since the PPP processing scheme is not changed except for the supplied ANTEX files, the results demonstrate the antenna phase variation impact on the tropospheric delay. They are shown in Table 6 and Figure 7. The results of the GA530 antenna are comparable to the performance of the LEIAR25.R3 and TRM55971.00 antennas. Supplying no ANTEX file at all, high ZTD biases in the order of −20 mm and −15 mm for the AV28 and ANN-MB-00 antennas are evident with the rectangular bracket. Using a circular plane reduced them to about 12 mm and 6.5 mm, respectively, while maintaining similar RMSE. When applying satellite PCO and PCV corrections, the boxplots in Figure 7 (Satellite PCO/PCV column) demonstrate that the standard deviation decreased notably. The presented offsets and consequently RMSE values limit the application of the data for meteorological purposes. After applying the elevation-only antenna corrections to the data, the bias and RMSE

decreased significantly for the antennas AV28 and ANN-MB-00. Remarkably, the applied corrections did not only increase the performance of the antennas with a circular plane, but also significantly decreased the error from using the rectangular bracket as base. Applying the azimuth-elevation dependent corrections to the data slightly increased the performance of the low-cost antennas. The bias increased slightly for the GA530 antenna and is presumably caused by the short observation time and outliers present in the residuals. The results are encouraging and demonstrate that the applied corrections for low-cost antennas make the data more attractive for tropospheric analysis. Particularly, the performance of the patch antenna ANN-MB-00 with a circular ground plane after correcting the phase pattern yields comparable results (0.47 mm bias and 4.02 mm RMSE) to those from geodetic quality antennas. To confirm these findings, longer observation times are recommended. The utilized IGS I14.ATX ANTEX file consisting of the elevation-only calibrations (without PCO estimations) for the antennas LEIAR25.R3 LEIT, TRM55971.00 NONE, Trimble GA530 and Trimble AV28 and U-blox ANN-MB-00 using a circular plane is attached as Supplementary Material.

4.3. Dual- versus Single-Frequency

The ionosphere-free linear combination, which is applied to dual-frequency data in PPP to eliminate the ionospheric delay, almost triples the noise in the data. Considering that noise in the data, including the noise in the antenna calibrations, is the major contributor to the bias and RMSE in the ZTD estimation, it can be useful to use L1 data only. Also, although this could change in the future, single-frequency receivers are more readily available than cheap dual-frequency receivers. By applying the SEID algorithm to L1-only observations, an artificial second frequency is generated from the L1 data using the interpolated ionospheric delay from a network of dual-frequency reference receivers. We apply this method to investigate if L1-only measurements may be favorable for ZTD estimations. The results are shown in Table 7. Using the original L1-only data and IGS ANTEX file, the SEID results indicate a similar performance for the high-quality antennas LEIAR25.R3, TRM55971.00 and GA530 compared to the L1 + L2 processing using the original IGS ANTEX file (Tables 5 and 6). Compared to results obtained using only satellite PCO/PCV corrections, the performance of the AV28 and ANN-MB-00 antennas increased significantly after applying SEID on the original data. The lowest bias (−0.40 mm) and RMSE (4.23) is observed with the ANN-MB-00 using a circular plane as base. The results with rectangular brackets demonstrate a ZTD RMSE of about 7 mm and a bias of −5 to −6 mm. Especially the bias of the datasets with rectangular brackets could be reduced to −0.36 mm (ANN-MB-00) and 2.26 mm (AV28) after applying elevation corrections to the L1 data. Applying the azimuth-elevation dependent corrections to the L1 input data resulted in minor improvements for the AV28 data and slight degradation of the ANN-MB-00 data.

These results suggests that the original L1 + L2 data is considerably affected by noise amplified by the ionosphere-free linear combination. Provided that a network of high quality surrounding stations exists, the approach demonstrates the strong advantage of the SEID algorithm to generate L2, which drastically reduces the noise on the target receiver data. There is, however, a trade-off between smoothing the true signal and the underlying noise. Utilizing this method almost eliminates the ZTD bias that was present in the data and the RMSE decreased to a level that is comparable to high-quality measurements which makes it attractive for meteorological applications such as water vapor estimation.

5. Conclusions

In this paper we investigated the PPP ZTD performance of a recently introduced low-cost dual-frequency receiver (U-blox ZED F9P) in combination with different antennas, ranging from geodetic to mass-market devices, with and without applying relative antenna calibrations. The conducted experiments demonstrated that the U-blox ZED-F9P dual-frequency receiver is very well capable to produce high-quality results, with the limiting factor being the quality of the receiving antenna. However, our results show that, using a simple-to-apply method to correct for the PCV

of cost-efficient receiver antennas, high quality results are achievable even for low-cost antenna. This is demonstrated by a field experiment, resulting in ZTD estimations of similar quality as with high-grade antennas.

The phase residuals with different antenna types were analyzed over a short baseline. The aim was to do a relative antenna calibration. The absolute antenna calibration pattern can then be computed from the absolute antenna calibration of the reference antenna and the relative calibration result from the short baseline experiment. For two of the tested antennas, the AV28 and ANN-MB-00 antennas, when used with a rectangular bracket as mounting point, very prominent elevation-based patterns were found (see Figure 5). These cases result in the highest RMSE phase residuals when compared to other antennas. Figure 6 indicates azimuth dependent patterns with strongly fluctuating amplitudes. The rectangular bracket presumably caused high residuals for L2 close to the horizon, but also in East and West direction for L1. Smaller residuals were obtained after repeating the experiment with a circular ground plane. These results showed that our approach is working, regardless of the size of the residual patterns. This suggests that our approach is feasible for even more challenging, multipath-prone environments. Additional uncertainty may be introduced by the smoothing technique we used to obtain PCVs in the 5-degree bin size required by the ANTEX standard. A lower binning size may further improve the results. Considering that only three days of data were used for the calibration, and that the antenna was not rotated, many of the azimuth-elevation bins were without data or had only few observations. For this reason the elevation-only based calibration is preferred over the azimuth-elevation based calibration.

The impact of the different antenna PCV corrections on PPP ZTD estimations has been analyzed for the tested antennas. Our results confirm that antenna pattern corrections are essential for PPP ZTD estimations. Applying satellite PCO/PCV corrections significantly decreases the standard deviation in the ZTD error compared to using no ANTEX corrections at all. Without applying receiver antenna corrections, the ANN-MB-00 and Trimble AV28 antennas with a rectangular bracket, resulted in a ZTD bias between -20 and -24 mm and similar standard deviations. With a circular plane the effect could be partially mitigated, but biases in the order of about 9.5 mm and 5.3 mm remained. The results suggest that a phase pattern is present for the low-cost antennas which we address by applying a relative antenna calibration. Applying elevation or azimuth-elevation dependent corrections to the data reduced the ZTD bias significantly and lowered the standard deviation. For example, when using the azimuth-elevation dependent corrections on the ANN-MB-00 antenna, the bias in the ZTD was reduced to -0.52 mm and an RMSE to only 3.77 mm. Results for the other antennas, and using elevation only patterns, were similar. This shows that the ZTD estimations achieve an error level that is comparable to high-grade antennas. Though the biases for ANN-MB-00 (rectangular bracket), AV28 (circular plane) and AV28 (rectangular bracket) could not be completely removed, they were reduced significantly to a level that makes the observations useful for tropospheric analysis.

The ionosphere-free linear combination used by the PPP solutions is very noisy. The noise in the L1 and L2 data, including the errors in the relative calibration, is basically tripled. For this reason we also looked at only using the L1 data in combination with the SEID algorithm to generate L2 data from an existing network of geodetic-grade receivers. In case no PCV correction is done for the receiving antenna, the bias and standard deviation in ZTD for the GA530, AV28 and ANN-MB-00 antennas, were smaller using L1 with SEID generated L2 data, than for the original dual-frequency data. This is a clear indication that the ionosphere-free linear combination on the original L1 + L2 data is considerably amplifying the noise present in the datasets. Using SEID in combination with the elevation or azimuth-elevation based L1 corrected data removed the biases almost entirely. The results are of comparable quality to ZTD estimations derived from the dual-frequency results.

This experiment uses exclusively GPS observations. As many of the low-cost receivers can track multiple GNSS systems, expanding the antenna calibrations to include multi-GNSS may further increase the application of the presented approach. Another interesting experiment would be to perform an absolute antenna calibration by a specialized company on low-cost antennas and compare

their calibrations to our results. Further work is needed to investigate, if the observed pattern is reproduced (or differs) from other antennas of the same model, so that the observed elevation-based phase pattern of an antenna can be applied to other antennas of the same model. This will be subject of future experiments.

Abbreviations

The following abbreviations are used in this manuscript:

ANTEX	Antenna Exchange Format
AP	Application Program
APC	Antenna Phase Center
AR	Ambiguity Resolution
CLI	Command-line User Interface
DD	Double Differences
DOY	Day-Of-Year
DPGA	Dutch Permanent GNSS Array
EKF	Extended Kalman Filter
ENU	East North Up
GMF	Global Mapping Function
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPT	Global Pressure/Temperature
GUI	Graphical User Interface
IGS	International GNSS Service
LEIAR25.R3	Leica AR25.R3
NMi	Nederlands Meetinstituut
PCO	Phase Center Offset
PCV	Phase Center Variation
PPP	Precise Point Positioning
PWV	Precipitable Water Vapor
RINEX	Receiver Independent Exchange Format
RMSE	Root Mean Square Error
SEID	Satellite-specific and Epoch-differenced Ionospheric Delay
TRM55971.00	Trimble Zephyr 2 Geodetic
USNO	US Naval Observatory
ZTD	Zenith Tropospheric Delay
ZWD	Zenith Wet Delay

Appendix A. Data Storing and Conversion

Data were stored in the binary U-blox logging format (*.ubx) on a local SD card of the Raspberry Pi. The logged files consisted of the UBX-RAWX messages which contained the raw measurements of the receiver (pseudorange, carrierphase, doppler and SNR on all recorded frequencies and satellite systems). The sampling interval was 1 Hz. Data were stored in 24 h batches of .ubx log files. For further processing, the data were converted to daily, 1Hz RINEX3 files. This step was performed with a combination of the RTKLIB application convbin (v. 2.4.3) and gfzrnx (v. 1.11) [36]. Figure A1 represents the work flow of this process. Metadata, if available, were added to the generated data.
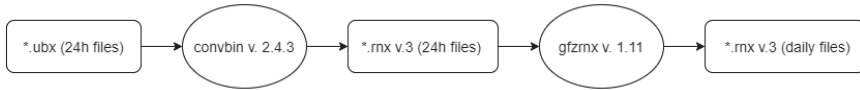
Figure A1. Sketch of *.ubx data format conversion to RINEX version 3.

Appendix B. Ambiguity Fixing Case Study

Since ambiguity fixing is essential to achieve highest accuracy results and RTKLIB offers several methods, a brief sensitivity study was conducted utilizing the continuous, instantaneous and fix-and-hold RTKLIB options with the same geodetic-grade antenna type on both, rover and base station (LEIAR25.R3 LEIT). For the sensitivity case study, one day of data from the highest quality antenna, LEIAR25.R3 LEIT was used. The continuous mode estimates phase biases over many epochs. It provides less outliers but when erroneous data are fed into the EKF it may remain over several epochs. The instantaneous method estimates integer ambiguities without constraining the previous (successful) fixes epoch-by-epoch. Fix-and-hold works similar to the Continuous AR, but tightly constrains the validated carrier-phase bias parameters in the next EKF update step. Further details can be obtained from the RTKLIB manual [37]. The results of the RTKLIB processing scheme are shown in Table A1. With 99.69% and 99.70% fixed solutions on L1 and L2 or only 2700 and 1614 float satellite observations on L1 and L2, the fix-and-hold solution performed best in our scenario.

Table A1. Fix and float satellite residuals summary of the LEIAR25.R3 LEIT antenna on DOY 79. For the fix-and-hold ambiguity solution, 'hold' epochs are considered as fix (Q==1).

AR-Method	Frequency	N (Epochs)	Fix N(%) Q==1	Float N(%) Q==2
Continuous	L1	873,729	835,273 (95.60)	38,456 (4.40)
Continuous	L2	531,297	508,782 (95.76)	22,515 (4.24)
Instantaneous	L1	873,729	302,696 (34.64)	571,033 (65.36)
Instantaneous	L2	531,297	189,191 (35.61)	342,106 (64.39)
Fix-and-hold	L1	873,729	871,029 (99.69)	2700 (0.31)
Fix-and-hold	L2	531,297	529,683 (99.70)	1614 (0.30)

References

1. Bevis, M.; Businger, S.; Herring, T.A.; Rocken, C.; Anthes, R.A.; Ware, R.H. GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. Geophys. Res. Atmos.* **1992**, *97*, 15787–15801. [CrossRef]

2. Bonafoni, S.; Biondi, R.; Brenot, H.; Anthes, R. Radio occultation and ground-based GNSS products for observing, understanding and predicting extreme events: A review. *Atmos. Res.* **2019**, *230*, 104624. [CrossRef]

3. Guerova, G.; Jones, J.; Douša, J.; Dick, G.; Haan, S.D.; Pottiaux, E.; Bock, O.; Pacione, R.; Elgered, G.; Vedel, H.; et al. Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. *Atmos. Meas. Tech.* **2016**, *9*, 5385–5406. [CrossRef]

4. Zumberge, J.; Heflin, M.; Jefferson, D.; Watkins, M.; Webb, F. Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J. Geophys. Res. B Solid Earth* **1997**, *102*, 5005–5017. [CrossRef]

5. Hauschild, A. Combinations of Observations. In *Springer Handbook of Global Navigation Satellite Systems*; Springer International Publishing: Cham, Switzerland, 2017; pp. 583–604._20. [CrossRef]

6. Barindelli, S.; Realini, E.; Venuti, G.; Fermi, A.; Gatti, A. Detection of water vapor time variations associated with heavy rain in northern Italy by geodetic and low-cost GNSS receivers. *Earth Planets Space* **2018**, *70*, 28. [CrossRef]

7. Krietemeyer, A.; ten Veldhuis, M.C.; van der Marel, H.; Realini, E.; van de Giesen, N. Potential of cost-efficient single frequency GNSS receivers for water vapor monitoring. *Remote Sens.* **2018**, *10*, 1493. [CrossRef]

8. Kouba, J.; Lahaye, F.; Tétreault, P. Precise Point Positioning. In *Springer Handbook of Global Navigation Satellite Systems*; Springer International Publishing: Cham, Switzerland, 2017; pp. 723–751._25. [CrossRef]

9. Mader, G.L. GPS antenna calibration at the National Geodetic Survey. *GPS Solut.* **1999**, *3*, 50–58. [CrossRef]

10. Gendt, G. *IGS Switch to Absolute Antenna Model and ITRF2005*; IGSMAIL-5438; IGS Central Bureau: Pasadena, CA, USA, 2006.

11. Görres, B.; Campbell, J.; Becker, M.; Siemes, M. Absolute calibration of GPS antennas: Laboratory results and comparison with field and robot techniques. *GPS Solut.* **2006**, *10*, 136–145. [CrossRef]

12. Schupler, B.R.; Allshouse, R.L.; Clark, T.A. Signal characteristics of GPS user antennas. *Navigation* **1994**, *41*, 276–296. [CrossRef]

13. Wübbena, G.; Schmitz, M.; Boettcher, G.; Schumann, C. Absolute GNSS antenna calibration with a robot: Repeatability of phase variations, calibration of GLONASS and determination of carrier-to-noise pattern. In Proceedings of the IGS Workshop, Darmstadt, Germany, 8–11 May 2006; pp. 8–12.

14. Geo++. GNSS Technologies. 2019. Available online: http://www.geopp.de/ (accessed on 23 September 2019).

15. IGS. ANTEX: The Antenna Exchange Format. 2019. Available online: http://www.igs.org/assets/txt/antex14.txt (accessed on 23 September 2019).

16. Araszkiewicz, A.; Völksen, C. The impact of the antenna phase center models on the coordinates in the EUREF Permanent Network. *GPS Solut.* **2017**, *21*, 747–757. [CrossRef]

17. Schmid, R.; Mader, G.; Herring, T. From relative to absolute antenna phase center corrections. In *Proceedings of the IGS Workshop and Symposium 2004*; Meindl, M., Ed.; Astronomical Institute, University of Bern: Bern, Switzerland, 2005; pp. 209–219.

18. Sidorov, D.; Teferle, F.N. Impact of antenna phase centre calibrations on position time series: preliminary results. In *IAG 150 Years*; Springer International Publishing: Cham, Switzerland, 2015; pp. 117–123.

19. Rothacher, M. Estimation of Station Heights with GPS. In *Vertical Reference Systems*; Drewes, H., Dodson, A.H., Fortes, L.P.S., Sánchez, L., Sandoval, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 81–90.

20. Ejigu, Y.G.; Hunegnaw, A.; Abraha, K.E.; Teferle, F.N. Impact of GPS antenna phase center models on zenith wet delay and tropospheric gradients. *GPS Solut.* **2018**, *23*, 5. [CrossRef]

21. Bar-Sever, Y.E.; Kroger, P.M.; Borjesson, J.A. Estimating horizontal gradients of tropospheric path delay with a single GPS receiver. *J. Geophys. Res. Solid Earth* **1998**, *103*, 5019–5035. [CrossRef]

22. Pacione, R.; Araszkiewicz, A.; Brockmann, E.; Dousa, J. EPN-Repro2: A reference GNSS tropospheric data set over Europe. *Atmos. Meas. Tech.* **2017**, *10*, 1689–1705. [CrossRef]

23. Loh, R.; Wullschleger, V.; Elrod, B.; Lage, M.; Haas, F. The US Wide-Area Augmentation System (WAAS). *Navigation* **1995**, *42*, 435–465. [CrossRef]

24. Noll, C.E. The Crustal Dynamics Data Information System: A resource to support scientific analysis using space geodesy. *Adv. Space Res.* **2010**, *45*, 1421–1440. [CrossRef]

25. Takasu, T. RTKLIB: An Open Source Program Package for GNSS Positioning. 2019. Available online: https://github.com/tomojitakasu/RTKLIB (accessed on 3 June 2019).

26. Takasu, T. RTKLIB Github Issue Tracker. 2019. Available online: https://github.com/tomojitakasu/RTKLIB/issues/457#issuecom\ment-473523905 (accessed on 3 June 2019).

27. Teunissen, P.J.; De Jonge, P.; Tiberius, C. Performance of the LAMBDA method for fast GPS ambiguity resolution. *Navigation* **1997**, *44*, 373–383. [CrossRef]

28. Araszkiewicz, A.; Kiliszek, D.; Podkowa, A. Height Variation Depending on the Source of Antenna Phase Centre Corrections: LEIAR25. R3 Case Study. *Sensors* **2019**, *19*, 4010. [CrossRef]

29. Herrera, A.M.; Suhandri, H.F.; Realini, E.; Reguzzoni, M.; de Lacy, M.C. goGPS: Open-source MATLAB software. *GPS Solut.* **2016**, *20*, 595–603. [CrossRef]

30. Lyard, F.; Lefevre, F.; Letellier, T.; Francis, O. Modelling the global ocean tides: Modern insights from FES2004. *Ocean Dyn.* **2006**, *56*, 394–415. [CrossRef]

31. Dow, J.; Neilan, R.; Rizos, C. The International GNSS Service in a changing landscape of Global Navigation Satellite Systems. *J. Geod.* **2009**, *83*, 191–198. [CrossRef]
32. Van Der Marel, H.; De Bakker, P. Single- versus Dual-Frequency Precise Point Positioning. *Inside GNSS* **2012**, *1*, 30–35.
33. Deng, Z.; Bender, M.; Dick, G.; Ge, M.; Wickert, J.; Ramatschi, M.; Zou, X. Retrieving tropospheric delays from GPS networks densified with single frequency receivers. *Geophys. Res. Lett.* **2009**, *36*. [CrossRef]
34. IGS. IGS Products. 2019. Available online: http://www.igs.org/products (accessed on 23 September 2019).
35. IGS. ANTEX Format Description. 2010. Available online: https://kb.igs.org/hc/en-us/articles/216104678-ANTEX-format-description (accessed on 24 April 2020).
36. Nischan, T. GFZRNX—RINEX GNSS Data Conversion and Manipulation Toolbox (Version 1.05). 2016. Available online: http://doi.org/10.5880/GFZ.1.1.2016.002 (accessed on 27 April 2020).
37. Takasu, T. RTKLIB ver. 2.4.2 Manual. 2013. Available online: http://www.rtklib.com/prog/manual_2.4.2.pdf (accessed on 3 June 2019).

Nine-Year Systematic Evaluation of the GPM and TRMM Precipitation Products in the Shuaishui River Basin in East-Central China

Xiaoying Yang [1], Yang Lu [1], Mou Leong Tan [2], Xiaogang Li [1], Guoqing Wang [3] and Ruimin He [3,*]

[1] Department of Environmental Science and Engineering, Fudan University, Shanghai 200433, China; xiaoying@fudan.edu.cn (X.Y.); 18210740057@fudan.edu.cn (Y.L.); 19110740037@fudan.edu.cn (X.L.)

[2] Geography Section, School of Humanities, Universiti Sains Malaysia, Penang 11800, Malaysia; mouleong@usm.my

[3] State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China; gqwang@nhri.cn

* Correspondence: rmhe@nhri.cn; Tel.: +86-21-65642192

Abstract: Owing to their advantages of wide coverage and high spatiotemporal resolution, satellite precipitation products (SPPs) have been increasingly used as surrogates for traditional ground observations. In this study, we have evaluated the accuracy of the latest five GPM IMERG V6 and TRMM 3B42 V7 precipitation products across the monthly, daily, and hourly scale in the hilly Shuaishui River Basin in East-Central China. For evaluation, a total of four continuous and three categorical metrics have been calculated based on SPP estimates and historical rainfall records at 13 stations over a period of 9 years from 2009 to 2017. One-way analysis of variance (ANOVA) and multiple posterior comparison tests are used to assess the significance of the difference in SPP rainfall estimates. Our evaluation results have revealed a wide-ranging performance among the SPPs in estimating rainfall at different time scales. Firstly, two post-time SPPs (IMERG_F and 3B42) perform considerably better in estimating monthly rainfall. Secondly, with IMERG_F performing the best, the GPM products generally produce better daily rainfall estimates than the TRMM products. Thirdly, with their correlation coefficients all falling below 0.6, neither GPM nor TRMM products could estimate hourly rainfall satisfactorily. In addition, topography tends to impose similar impact on the performance of SPPs across different time scales, with more estimation deviations at high altitude. In general, the post-time IMERG_F product may be considered as a reliable data source of monthly or daily rainfall in the study region. Effective bias-correction algorithms incorporating ground rainfall observations, however, are needed to further improve the hourly rainfall estimates of the SPPs to ensure the validity of their usage in real-world applications.

Keywords: satellite precipitation products; evaluation; daily rainfall; hourly rainfall; GPM; TRMM

1. Introduction

Precipitation is an important component of the hydrological cycle [1,2]. Accurate and high-resolution precipitation data are crucial in different fields such as weather forecast, disaster preparation and prevention, and water resource management [3,4]. The quality and resolution of precipitation inputs can also significantly affect the performance of various hydrological, climatic, and atmospheric models [5].

However, obtaining suitable precipitation data could be challenging for researchers as well as practitioners. Availability of traditional ground observations has been limited because of the inadequate and uneven distribution of rain gauges, especially in developing countries, mountainous and remote

areas, and over oceans [6]. On the other hand, although weather radar products can provide rainfall observations over a wide region [7], they are subjected to both random and systematic errors [8–11]. Random errors could arise from the sub-grid horizontal and vertical variability of rainfall and the noise of the radar hardware system, while systematic errors may originate from sources such as drifts in radar calibration constant, systematic variations in the reflectivity–rain-rate relationship, and strong gradients in the reflectivity profile [12]. The presence of complex topography may further amplify some of the error sources [13].

In recent years, with the rapid development of remote sensing techniques, satellite precipitation products (SPPs) have been increasingly applied in monitoring precipitation patterns [14–16]. Deriving precipitation products through satellite remote sensing has the advantages of wide coverage and high spatiotemporal resolution, which complement traditional ground gauge measurements. For example, the Tropical Rainfall Measuring Mission (TRMM) satellite launched in 1997 has been extensively used in hydrological modelling and climate change studies. Li et al. [17] found overall good linear relationships between TRMM and ground rainfall observations at both daily and monthly time steps in the Xinjiang catchment, China. Bharti and Singh [18] compared TRMM 3B42V7 with the gauge-based measurements at different altitudes in the northwest Himalayan region. They found that the satellite performed satisfactorily in the altitude range of 1000–2000 m, but poorly over higher-altitude regions at a daily time step.

In 2014, the National Aeronautics and Space Administration (NASA) of U.S. and JAXA (Japan Aerospace Exploration Agency) jointly developed a new generation of Global Precipitation Measurement (GPM) satellites. In addition to inheriting the advantages of the TRMM satellites in detecting precipitation in the tropics, GPM satellites provide global precipitation estimates for a wider quasi-global coverage (60° N–60° S) at a much higher spatiotemporal resolution ($0.1° \times 0.1°$ and 30-min interval). Much research has concluded that GPM products have improved in terms of both rainfall observation accuracy and hydrological simulation performance compared to TRMM products. For example, Tan et al. [19] and Sharifi et al. [20] compared the accuracy of rainfall observations between IMERG (integrated multi-satellite retrievals for GPM) and TRMM in Singapore and India, respectively. In both studies, all evaluation indices had indicated a better performance of IMERG than TRMM in providing monthly and daily rainfall data. In China, Tang et al. [21] analyzed the errors of IMERG and TRMM products in six sub-regions of Mainland China and found that IMERG had improved the accuracy of precipitation observations in the mid-high latitude as well as arid regions. In addition, they observed that IMERG could better reproduce the probability density function of rainfall, especially in the range of lower rainfall intensity.

In June 2019, the IMERG product was upgraded from Version 5 (V5) to Version 6 (V6) by reducing biases based on the new Global Precipitation Climatology Centre (GPCC) monthly precipitation records. Meanwhile, TRMM data before 2014 have also been reprocessed with the latest algorithm of the IMERG V6. Until now, few studies have been carried out to evaluate the performance of the latest IMERG and TRMM products.

Furthermore, the majority of previous research has evaluated SPP products at the monthly or daily scale, although both GPM and TRMM products contain hourly rainfall products. High quality hourly rainfall data have been found to be valuable to various hydrological applications around the world [22]. For example, Zhou and Wu [23] found that the precipitation intensity and distribution characteristics of typhoons in China could be better analyzed with hourly precipitation than daily observations at automatic weather stations. Yang et al. [24,25] found that the SWAT (Soil and Water Assessment Tool) model built on hourly rainfall could yield much better performance in simulating daily streamflow and monthly nutrient loads than the SWAT model built on daily rainfall in the Upper Huai River basin of China. Boithias et al. [26] found that the SWAT model built on hourly rainfall could better predict discharge over long periods of time than the MARINE model in the Mediterranean coastal Têt River basin (Southwestern France).

Compared to daily rainfall, hourly rainfall data are much more difficult to obtain because of several reasons. Firstly, much fewer gauges can or will record the amount of rainfall at an hourly or finer interval worldwide. Secondly, hourly rainfall data are generally not free to the public. Purchase of hourly rainfall data might be too expensive for researchers or practitioners in some regional studies. Finally, authorities in some regions may consider hourly rainfall records as sensitive data, thus denying their access to the public citing security reasons. In view of the limited access to hourly rainfall data globally, SPPs may provide a much-needed alternative for deriving such products. So far, few studies have been carried out to evaluate the capability of SPPs in providing hourly rainfall estimates.

To fill in the gaps, this study aims to evaluate the accuracy of the latest GPM and TRMM rainfall products across the monthly, daily, and hourly scales based on the ground rain gauge measurements between January 2009 and December 2017 in the Shuaishui River Basin (SRB) of eastern Central China. The Shuaishui River is the headwater tributary to the Qiantangjiang River, the main river flowing across the Zhejiang Province of China. With water quality inferior to the Class III standard at 50.5% of its total river length, the Qiantangjiang River Basin is faced with severe water security concerns [27,28]. As the critical ecological barrier to the Qiantangjiang River, the hydrological conditions of the SRB has direct impact on the downstream ecological environment.

Essentially a hilly watershed, SRB is characterized with complex terrains and obvious vertical height difference. Precipitation in the basin is abundant, but also highly seasonal. Steep slopes combined with ample rainfall in summer have aggravated the risk of natural disasters such as floods and mudslides [29]. The flood in June 2016 in the SRB, for instance, has affected 58,000 people with a direct economic loss of 168 million RMB. SRB, therefore, presents an ideal referencing region for evaluating the suitability of using SPPs in the sub-tropical hilly regions with large inter-annual and intra-annual rainfall variabilities.

2. Study Area

Approximately 159 km in length, the Shuaishui River originates from the Hutou mountain ranges and flows across the Xiuning County before pouring into the Xinanjiang River at the Tunxi district of Huangshan City. The SRB (117°39′–119°26′ E and 29°24′–31°1′ N) has a total area of 1522 km^2 (Figure 1). Dominated by a hilly terrain, more than 70% of the basin is at an altitude of above 500 m. Land use and land cover in the basin is mainly forestland and cultivated land, which respectively accounts for 78.9% and 14.6% of the total coverage.
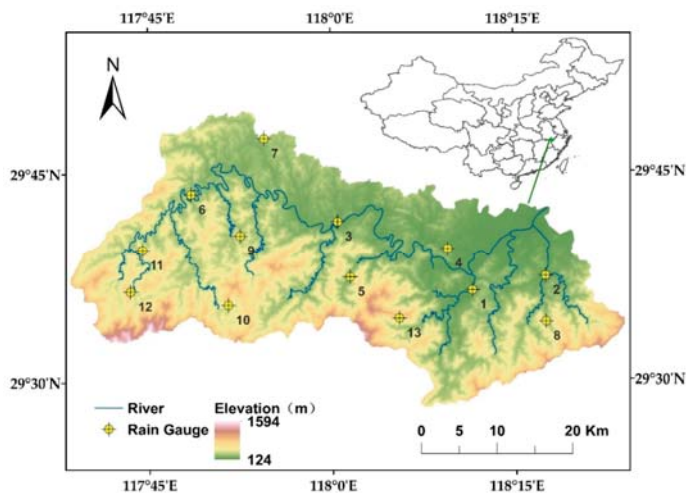

Figure 1. Map of Shuaishui River Basin.

Located in the subtropical monsoon climate zone, rainfall in the SRB is usually abundant. Between 2009 and 2017, mean annual rainfall observed by rain gauges in the basin ranges from 1747 mm in 2013 to 2700 mm in 2015 with an overall average of 2278 mm (Figure A1). Within each year, mean monthly rainfall usually increases steadily from January to May and peaks in June. Precipitation in June alone could account for more than one-fifth of annual total rainfall. After June, monthly rainfall falls sharply and exhibits an overall decreasing trend till the end of the year (Figure A2).

3. Materials and Methods

3.1. Satellite Precipitation Products

The TRMM satellite was launched in 1997 through a joint space mission between the NASA of U.S. and the National Space Development Agency of Japan [30]. TRMM carries five instruments, including a suite of three rainfall sensors (Precipitation Radar (PR), TRMM Microwave Imager (TMI), Visible and Infrared Sensor (VIRS)) and two related instruments (Lightening Imaging Sensor (LIS) and Clouds and the Earth's Radiant Energy System (CERES)). The TRMM Multi-satellite Precipitation Analysis (TMPA) products combine infrared (IR) data from geostationary satellites, such as GOES-W, GOES-E, GMS, Meteosat-5, Meteosat-7, and NOAA-12, with microwave (MW) data from multiple satellites including TMI/TRMM (TRMM Microwave Imager), SSMI/DMSP (Special Sensor Microwave Imager/Defense Meteorological Satellite Program), AMSU/NOAA (Advanced Microwave Sounding Unit/National Oceanic and Atmospheric Administration), and AMSR-E (Advanced Microwave Scanning Radiometer-EOS) [31]. The TMPA products are produced in the following four stages. First, the MW precipitation estimates are calibrated and combined using algorithms such as sensor-specific versions of the Goddard Profiling Algorithm (GPROF). Secondly, IR precipitation estimates are created using the calibrated MW precipitation. Thirdly, the MW and IR precipitation estimates are combined. Finally, rain gauge data are incorporated. Detailed descriptions of the algorithms and steps for producing the TMPA products could be found in Huffman et al. (2007) [32] and Huffman et al. (2018) [33].

In May 2012, the TMPA was upgraded from version 6 (V6) to version 7 (V7) by implementing the latest version of re-calibration algorithm and using the new GPCC monthly precipitation products for bias correction. The TMPA 3B42 consists of two products: the near-real-time product (3B42RT) and the post-processed product (3B42). The 3B42RT product, which is released approximately 9 h after real-time, spans the latitude belt from 50° N to 50° S. In contrast, with a more extensive coverage from 60° N to 60° S, the 3B42 product is released 10–15 days after each month when the bias correction has been made based on ground gauge records.

As a global successor of TRMM, the GPM project is launched in 2014 to provide global precipitation observations. The GPM satellite is equipped with an advanced Dual-frequency Precipitation Radar (DRP) that observes the internal structure of storms within and under the clouds, and a GPM Microwave Imager (GMI) that measures the type, size, and intensity of precipitation. The DPR is more sensitive than its TRMM predecessor especially in the measurement of light rainfall and snowfall in high latitude regions.

In March 2014, NASA released its first GPM-era global precipitation product–IMERG (Integrated Multi-satellites Retrievals for GPM). The IMERG algorithm is designed to inter-calibrate, interpolate, and merge all available satellite MW precipitation estimates, MW-calibrated IR satellite estimates, gauge measurements, and other potential precipitation estimates at fine spatial and temporal resolution worldwide. Its inter-calibration of available MW data is similar to TMPA, but further interpolated and re-calibrated by the Climate Prediction Center (CPC) morphing Kalman Filter technique and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks—Cloud Classification System (PERSIANN-CCS) [34,35].

IMERG includes three products with different latencies: the near-real-time 'Early' (near real-time with a latency of 6 h) run (IMERG-E), the near-real-time 'Late' (reprocessed near real-time with a

latency of 18 h) run (IMERG-L), and the post-real-time 'Final' (gauge-adjusted with a latency of four months) run (IMERG-F). The algorithm for the IMERG was upgraded from Version 5 (V5) to Version 6 (V6) to reduce bias and improve consistency among different IMERG runs in June 2019. For example, the 'displacement vectors' in V6 are computed using the Modern Era Retrospective Reanalysis 2 (MERRA-2) and Goddard Earth Observing System (GEOS) model Forward Processing (FP) data instead of the previously used infrared data, which helps ensure consistency in the vectors between the Final Run and the Early and Late Runs.

In this study, we aim to evaluate the performance of a total of five SPPs in the SRB, including the Early, Late, and Final runs of the IMERG V6 products (0.1° × 0.1° and 30-min interval), and the near-real-time and post-processed runs of the TMPA V7 products (0.25° × 0.25° and 3-hour interval). The SPPs datasets are all downloaded from the NASA website (https://disc.gsfc.nasa.gov/). After being downloaded, the SPPs datasets are adjusted to the local time, which is eight hours ahead of Coordinated Universal Time (UTC). Since the IMERG V6 and TMPA V7 products respectively contain 30-min and 3-h rainfall estimates, they need to be processed before being evaluated at different temporal scales. At the daily and monthly scale, both IMERG and TMPA data are directly aggregated to the corresponding levels for comparison with ground measurements. For evaluation at the hourly scale, the TMPA hourly rainfall estimates are obtained by assuming a constant rainfall intensity over the 3-hour period.

3.2. Ground Rainfall Measurements

Hourly and daily precipitation records from 2009 to 2017 at a total of 13 rainfall stations in the SRB (Figure 1) are obtained from the hydrological yearbook series published by Ministry of Water Resources of China. The rain gauges used at the rainfall stations are tipping buckets. All of the rainfall data have gone through strict quality control following the relevant China's industry standards such as QX/T 118-2010 (quality control of surface meteorological observation data) before being published. There are no rainfall data missing at the 13 stations. In the SRB, daily rainfall has been recorded throughout the year, while hourly rainfall only documented for the relatively wet period from April to October. Correspondingly, the monthly and daily estimates of the SPPs are evaluated throughout the year, while their hourly estimates are only assessed within the seven months.

3.3. Evaluation Metrics

A total of four continuous metrics are used to evaluate the quality of satellite precipitation products in the SRB. Correlation coefficient (CC) is used to quantify the linear correlation between satellite precipitation estimates and ground measurements; it varies between −1 and 1, with a value close to 0 indicating little correlation. Root-mean-square error (RMSE) quantifies the degree of dispersion between satellite precipitation and measured precipitation, which can reflect the overall error level and accuracy of SPPs [36]; mean absolute difference (MAD) evaluates the magnitude of the average difference between satellite precipitation and measured precipitation. Smaller values of RMSE and MAD indicate a better performance of the SPPs. Relative bias (RB) measures the systematic bias of satellite precipitation compared with gauge observations. A positive and negative RB indicates overestimation and underestimation, respectively. As a rule of thumb, SPPs can be considered as reliable when RB falls between −10% and 10% and CC exceeds 0.7 [37].

The four continuous metrics are calculated as [38–40]

$$CC = \frac{\sum_{i=1}^{n} (X_i^o - \overline{X}^o)(X_i^s - \overline{X}^s)}{\sqrt{\sum_{i=1}^{n} (X_i^o - \overline{X}^o)^2} \sqrt{\sum_{i=1}^{n} (X_i^s - \overline{X}^s)^2}} \quad (1)$$

$$RB = \frac{\sum_{i=1}^{n} (X_i^s - X_i^o)}{\sum_{i=1}^{n} X_i^o} \times 100\% \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i^s - X_i^o)^2}{n}} \tag{3}$$

$$MAD = \frac{\sum_{i=1}^{n}|X_i^o - X_i^s|}{n} \tag{4}$$

where n is the number of the simulated and observed data pairs; X_i^s and X_i^o denote the ith simulated and observed amount, respectively; $\overline{X}^s$ and $\overline{X}^o$ are the mean of the simulated and observed data, respectively.

Besides the continuous metrics, three categorical evaluation metrics are used to evaluate the precipitation detection capability of the SPPs, which include probability of detection (POD), false alarm rate (FAR), and critical success index (CSI). POD represents the ratio of correctly detected precipitation occurrences by the SPPs to the total number of actual precipitation occurrences. With an optimal value of 1, a higher POD indicates that the SPP is more capable of detecting the actual precipitation occurrences. FAR calculates the ratio of falsely detected precipitation occurrences to the total number of detected precipitation occurrences. With an optimal value of 0, a lower FAR indicates that the SPP is less likely to yield false precipitation occurrences. CSI incorporates both missed events and false detections in its calculation [41]. With an optimal value of 1, a higher CSI indicates a better performance of the SPP with more correct detections as well as fewer false alarms of precipitation occurrences. Based on the number of hits (H), false alarms (F), and misses (M) (Table 1), the three categorical metrics are calculated as

$$POD = \frac{H}{H+M} \tag{5}$$

$$FAR = \frac{F}{H+F}, \tag{6}$$

$$CSI = \frac{H}{H+M+F} \tag{7}$$

where S represents rain gauge observation; P represents satellite rainfall estimate; H (hits) represents the number of cases when both the rain gauge and the satellite determine the rainfall to equal or exceed the threshold; F (false alarms) represents the number of cases when the satellite determines the rainfall to equal or exceed the threshold but not the rain gauge; M (misses) represents the number of cases when the rain gauge determines the rainfall to equal or exceed the threshold but not the satellite; and Z (correct negatives) represents the number of cases when both the rain gauge and the satellite determine the rainfall to fall below the threshold.

Table 1. Contingency table between rain gauge observations and satellite precipitation estimates.

SPPs Estimates	Rain Gauge Observations	
	$S \geq$ Threshold	$S <$ Threshold
$P \geq$ Threshold	H	F
$P <$ Threshold	M	Z

3.4. Analysis of Variance (ANOVA)

To evaluate the rainfall estimation performance of the five satellite precipitation products, four continuous metrics (*CC, RB, RMSE,* and *MAD*) are respectively calculated at each of the 13 rainfall stations across the monthly, daily, and hourly scales. Previous studies have mostly used the mean values of the metrics over all rainfall stations to compare the rainfall estimation performance among the SPPs. This simple averaging approach, however, does not account for the variability in the metrics among the rainfall stations. Furthermore, it is incapable of determining the significance of the difference between the SPPs.

In view of the deficiency, we adopt the one-way analysis of variance (ANOVA) to statistically evaluate the difference in metrics between the SPPs. In the ANOVA, satellite precipitation product

type is used to designate the five groups of metrics for comparison. If the ANOVA determines there is some significant difference in the mean metrics among the SPPs, multiple commonly used posterior comparison tests—including the Bonferroni, Sidak, Tukey, and Scheffe tests built in the Origin 2018 Statistical Package—are further used to identify the pairs of SPPs whose mean metrics are indeed statistically different.

4. Results and Discussions

4.1. Evaluation at the Monthly Scale

4.1.1. Temporal Analysis

Figure 2 compares the mean of the observed monthly precipitation in the SRB with that of the five SPPs from January 2009 to December 2017. Figure A3 compares the scatterplots between observed and estimated monthly rainfall among the SPPs. In general, all five SPPs are capable of capturing the overall trend of monthly precipitation variations. The annual CCs of the five SPPs are all above 0.85, while those of IMERG_F and 3B42 even exceed 0.95. The IMERG products all exhibit a tendency of underestimating rainfall, especially in wet months. In particular, for June, 2011 whose monthly rainfall reached as high as 1109 mm, IMERG_E, IMERG_L, and IMERG_F give a low estimate of 608, 608, and 710 mm, respectively.
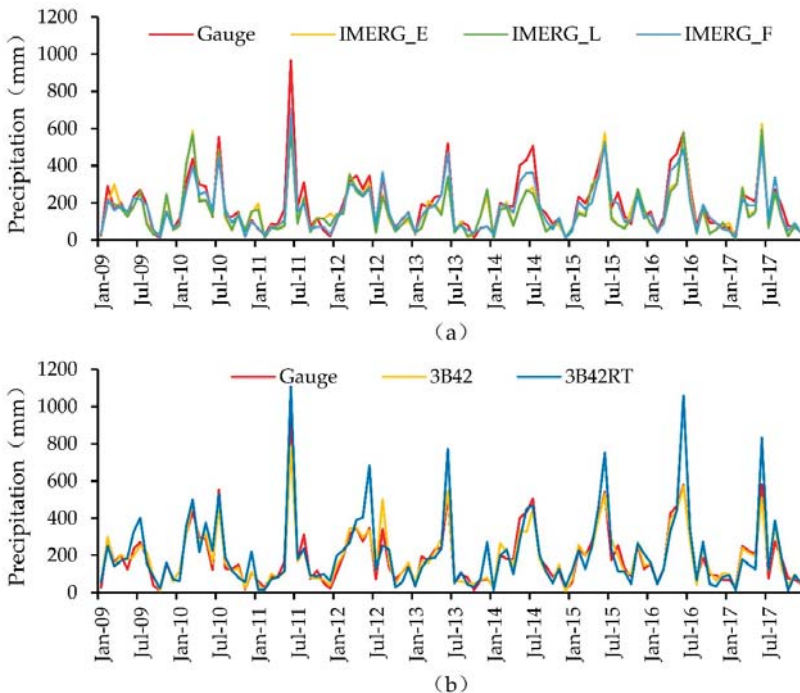
Figure 2. Comparison of mean monthly precipitation observations with the estimates of five SPPs from 2009 to 2017: (**a**) IMERG products; (**b**) TRMM products.

Table 2 compares the mean values of the four continuous evaluation metrics over the 13 rainfall stations among the five SPPs both annually and seasonally. With the highest RMSE of 101.42 mm and MAD of 65.72 mm, the 3B42RT product deviates the most from historical rainfall observations annually. A closer examination of the seasonal changes in RMSE and MAD, however, have shown that

its considerably larger deviation in summer is the main cause. In the three seasons other than summer, 3B42RT actually deviates less than the two near-real-time IMERG products (Table 2).

Table 2. Mean evaluation metrics of the SPPs at monthly scale.

Metrics	Temporal Scale	IMERG_E	IMERG_L	IMERG_F	3B42	3B42RT
CC	Annual	0.86	0.86	0.97	0.95	0.89
	Spring [a]	0.72	0.77	0.95	0.93	0.82
	Summer [a]	0.89	0.90	0.95	0.93	0.89
	Fall [a]	0.65	0.69	0.88	0.83	0.74
	Winter [a]	0.69	0.70	0.98	0.97	0.81
RMSE (mm)	Annual	86.69	87.64	53.82	54.13	101.42
	Spring	90.17	89.50	47.23	44.02	75.09
	Summer	118.98	121.71	87.19	87.36	170.08
	Fall	53.88	52.51	30.64	35.79	49.90
	Winter	67.13	68.93	25.18	26.41	61.58
RB (%)	Annual	−13.63	−16.80	−10.51	0.61	10.69
	Spring	−13.74	−19.33	−12.00	−0.02	−4.91
	Summer	−19.13	−22.71	−9.85	−1.48	25.77
	Fall	−11.20	−9.84	−6.35	0.12	−3.89
	Winter	0.79	0.15	−11.88	9.68	19.06
MAD (mm)	Annual	61.59	61.77	35.79	37.07	65.72
	Spring	71.80	70.72	37.99	34.34	60.64
	Summer	84.78	87.90	65.17	65.30	118.77
	Fall	43.50	41.97	23.17	28.40	40.75
	Winter	46.28	46.48	16.83	20.25	42.74

[a] Spring extends from March to May; Summer extends from June to August; Fall extends from September to November; Winter extends from December to the following February.

Annually, the three IMERG products tend to underestimate monthly precipitation, while the two TMPA products behave the opposite. Seasonally, IMERG_F tend to underestimate monthly precipitation throughout the year, while the other two IMERG products act the same except that they tend to slightly overestimate in winter. In contrast, the RB of 3B42 remains close to zero all over the year except it approaches 10% in winter. Meanwhile, the RB of 3B42RT fluctuates much more ranging from -4.9% in spring to 25.8% in summer (Table 2). In terms of the other three continuous metrics, the five SPPs have exhibited somewhat similar seasonal patterns of change. For example, the CCs of the SPPs all reach or approach their peak values in summer, while decreasing to the bottom in fall. Meanwhile, both the RMSEs and MADs of the SPPs all rise to the top in summer, drop to medium in spring, and down to the lowest in fall and winter (Table 2). The seasonal changes in RMSEs and MADs correspond closely to the changes in the magnitude of seasonal rainfall.

Except for RB, the two post-time products (IMERG_F and 3B42) perform significantly better than the rest real-time or near real-time products both annually and seasonally, with a noticeably higher value of CC (e.g., 0.97 and 0.95 annually) as well as considerably lower values of RMSE (e.g., 53.82 and 54.13 mm annually), and MAD (e.g., 35.79 and 37.07 mm annually) (Table 2). The findings of the overall better performance of the two post-time SPPs products compared to the real-time or near real-time products at the monthly time scale are not surprising, since both are generated after the adjustment of real-time products based on monthly measurements of ground rain gauges [42], although which may not include the 13 rain gauges covered in our study. With their annual CC and RB values exceeding the good performance thresholds, both IMERG_F and 3B42 can be regarded as reliable sources of monthly precipitation in the SRB. Similar to our study, previous studies have also observed satisfactory performance of IMERG_F and 3B42 in monthly rainfall estimation [43,44].

4.1.2. Spatial Variation

Figure 3 compares the spatial distribution of the four annual continuous evaluation metrics among the SPPs at the monthly scale. The spatial distributions of the CCs varies considerably among the IMERG products, while staying similar between the TMPA products. Topography does not seem to be a significant influencing factor of the CCs, although some stations at higher altitude (station 12 and 13) do have lower CCs in all five SPPs. In terms of RMSEs and MADs; however, topography plays a more eminent role. For all five SPPs, both metrics tend to get larger at higher altitude. In particular, RMSEs and MADs of three stations at high altitude (station 10, 12, and 13) consistently surpass those of the rest stations. In addition, the RBs exhibit a similar spatial pattern among the IMERG products. They all tend to underestimate monthly rainfall more severely at high altitude (e.g., stations 12 and 13). However, for the TMPA products, no clear pattern in the spatial distribution of RB could be observed.
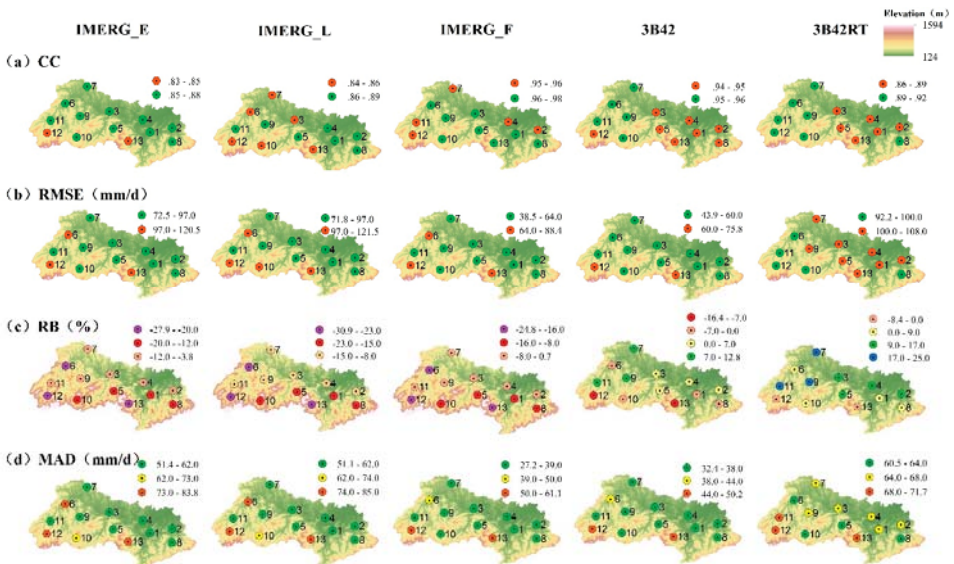
Figure 3. Spatial distribution of the annual continuous evaluation metrics at monthly scale: (**a**) *CC*; (**b**) *RMSE*; (**c**) *RB*; and (**d**) *MAD*.

Similar to our study, Milewski et al. [45] also found that elevation was a key factor affecting the accuracy of the TMPA products in Northern Morocco. The CCs of all four TMPA products at the low elevation class (0–500 m) consistently surpassed those at the medium (500–1000 m) and high (>1000 m) elevation classes. Contrary to the CCs, the normalized RMSEs at the low elevation class were consistently smaller.

4.2. Evaluation at the Daily Scale

Figure A4 compares the scatterplots between observed and estimated daily rainfall among the SPPs. Daily rainfall estimates by the IMERG_F cluster the closest around the 1:1 line, while 3B42RT estimates deviating the largest with a strong tendency of overestimation. For the daily scale assessment of SPPs, four continuous (CC, RMSE, RB, and MAD) and three categorical (POD, CSI, FAR) evaluation metrics are calculated using daily rainfall records at the 13 rainfall stations and the corresponding SPP estimates from 1 January 2009 to 31 December 2017.

4.2.1. Continuous Evaluation Metrics

(1) Temporal Variation

Table 3 compares the mean values of the continuous evaluation metrics among the SPPs at the daily scale. Annually, all four continuous evaluation metrics except RB have indicated a relatively better performance by the IMERG products in estimating daily rainfall. For example, the annual RMSEs of the IMERG products range from 9.66 mm/d (IMERG_F) to 11.30 mm/d (IMERG_E), while those of the TMPA products both exceed 11.50 mm/d. Seasonally, however, while IMERG_F generally remains the best product for estimating daily rainfall, the other two near-real-time IMERG products tend to perform better than the TMPA products in spring and summer, but often worse in fall and winter (except for RB).

Table 3. Mean continuous evaluation metrics of the SPPs at daily scale.

Metrics	Temporal Scale	IMERG_E	IMERG_L	IMERG_F	3B42	3B42RT
CC	Annual	0.73	0.75	0.81	0.75	0.75
	Spring [a]	0.71	0.77	0.79	0.74	0.73
	Summer [a]	0.80	0.81	0.83	0.78	0.77
	Fall [a]	0.64	0.65	0.73	0.70	0.68
	Winter [a]	0.66	0.66	0.82	0.72	0.68
RMSE (mm)	Annual	11.30	11.07	9.66	11.54	12.96
	Spring	12.23	11.04	10.61	13.20	13.05
	Summer	14.52	14.15	13.74	15.04	18.55
	Fall	8.03	8.56	6.67	7.56	7.82
	Winter	9.03	9.48	5.36	8.43	9.66
RB (%)	Annual	−13.15	−16.33	−9.99	1.21	11.36
	Spring	−11.89	−17.62	−11.55	0.47	−4.43
	Summer	−16.96	−20.62	−7.51	−0.59	26.94
	Fall	−9.68	−8.06	−4.31	0.56	-3.47
	Winter	2.98	2.23	−9.88	9.57	18.94
MAD (mm)	Annual	4.52	4.19	4.00	4.77	5.21
	Spring	5.45	4.81	4.90	5.81	5.71
	Summer	6.72	6.27	6.37	7.12	8.53
	Fall	2.93	2.85	2.60	2.91	2.98
	Winter	2.98	2.85	2.19	3.19	3.55

[a] Spring extends from March to May; Summer extends from June to August; Fall extends from September to November; Winter extends from December to the following February.

Within each SPP family, the order of the daily rainfall estimation accuracy is largely consistent, i.e., IMERG_F > IMERG_L > IMERG_E (except for RB) and 3B42 > 3B42RT. Moreover, except for RB, the five SPPs have exhibited rather similar seasonal patterns of change in their daily metrics. For example, the CCs of the five SPPs all peak in summer and fall to the bottom in fall. The RMSEs and MADs of the SPPs all tend to peak in summer, decline in spring, and down to the lowest in fall and winter. Throughout the year, the Shuaishui River Basin is affected by different climatic systems. Mainly under the influence of the high-altitude trough, precipitation in fall and winter is mostly brought by stratiform clouds, which tends to be stable and therefore easier to measure. In spring and summer, however, the convective component in the precipitation system increases due to the Meiyu front and shear line system. Both thermal convection precipitation under the control of the Western Pacific Subtropical High (WPSH) and rainstorms caused by the typhoon system increase the difficulty of obtaining accurate measurement of rainfall because of their characteristics of short duration and high spatial heterogeneity. The differences in climatic systems have led to different seasonal rainfall characteristics, with more rainy days and higher rainfall intensity in spring and summer than in fall and winter. Between 2009 and 2017, there have been 529 and 589 rainy days in spring and summer,

compared to 425 and 452 days in fall and winter. Days with precipitation < 1 mm account for 44.2% and 41.8% in fall and winter, compared to 30.6% and 26.7% in spring and summer. Meanwhile, days with precipitation > 50 mm account for 4.9% and 8.3% in spring and summer, compared to around 1.5% in fall and winter. The CCs' peaking in summer could be attributed to the season's large variability in daily rainfall, whose general pattern of change is relatively well captured by the SPPs. However, because there are more days with heavy precipitation in summer, the absolute errors of SPP estimates remain the largest in the season. Similarly, the lowest RMSEs and MADs in fall and winter are probably owing to their dominance of days with lower precipitation.

Similar to our findings, Su et al. [46] concluded that the post-time IMERG-F product, with a CC of 0.79, RMSE of 6.31 mm/d, and RB of 9.04%, was the best IMERG product for estimating daily rainfall in the Upper Huai River Basin of China. Meanwhile, Anjum et al. [43] found the post-time 3B42 V7 product, with a CC of 0.70 and RB of 14.77%, performed better than the real time 3B42RT product for estimating daily rainfall in Pakistan.

Compared to those at the monthly scale, the CCs of all five SPPs have decreased considerably at the daily scale. For instance, the annual CC of IMERG_F drops from 0.97 at the monthly scale to 0.81, while the annual CCs of the other SPPs all drop further to around 0.75. In contrast to the CCs, the RBs of the SPPs at the daily scale are more similar to those at the monthly scale in terms of both their signs and magnitude. Annually, all three IMERG products tend to underestimate daily precipitation with the lowest RB of −9.99%, while both TMPA products tend to overestimate with the lowest RB of 1.21%. Seasonally, the IMERG family products tend to underestimate daily rainfall in all four seasons except the two near-real-time products in winter. In contrast, 3B42RT exhibits a strong tendency of overestimation in summer and winter, while 3B42 only in winter.

(2) Statistical Performance Comparison among the SPPs

Figure 4 compares the boxplots of the four annual continuous evaluation metrics among the five SPPs. For example, Figure 4a contains five boxplots, which respectively depict the distribution of annual CCs of the five SPPs across the 13 rainfall stations. At the daily scale, the annual CCs of the IMERG_F range from 0.79 to 0.82 among the 13 rainfall stations, compared to 0.71–0.75 for IMERG_E, 0.73–0.77 for IMERG_L and 3B42RT, and 0.74–0.77 for 3B42.

One-way ANOVA could be used to assess whether the mean values of the four continuous metrics are significantly different among the five SPPs. One critical pre-condition of performing ANOVA is to ensure the homogeneity of variance among the compared groups. In this study, we use the Levene's Test to compare the variance of the metrics among the five SPPs, which confirm that all four metrics could meet the requirement of homogeneity of variance. The subsequent one-way ANOVA has concluded that the mean values of all four metrics are significantly different among the SPPs at the significance level (α) of 0.05 (Figure 4).

In view of the significant ANOVA results, multiple posterior comparison tests—including the Bonferroni, Sidak, Tukey, and Scheffe tests—are further conducted to identify the pairs of SPPs whose mean metrics are truly significantly different. In Figure 4, two SPPs are connected with a black dotted line if posterior comparison tests have concluded a non-significant difference between their means at the α level of 0.05. As shown in the figure, the mean values of CC are significantly different between all pairs of SPPs except between IMERG_L and the two TMPA products as well as between the TMPA products themselves; the mean values of RMSE are all significantly different except between 3B42 and the two near real-time IMERG products as well as between the two near real-time IMERG products themselves; the mean values of RB are all significantly different except between the three pairs of IMERG products; the mean values of MAD are all significantly different except between IMERG_L and the other two IMERG products as well as between IMERG_E and 3B42. It is worth noting that the posterior comparison tests have shown that IMERG_F is the single IMERG product that is significantly different from the TMPA products in terms of all four metrics.
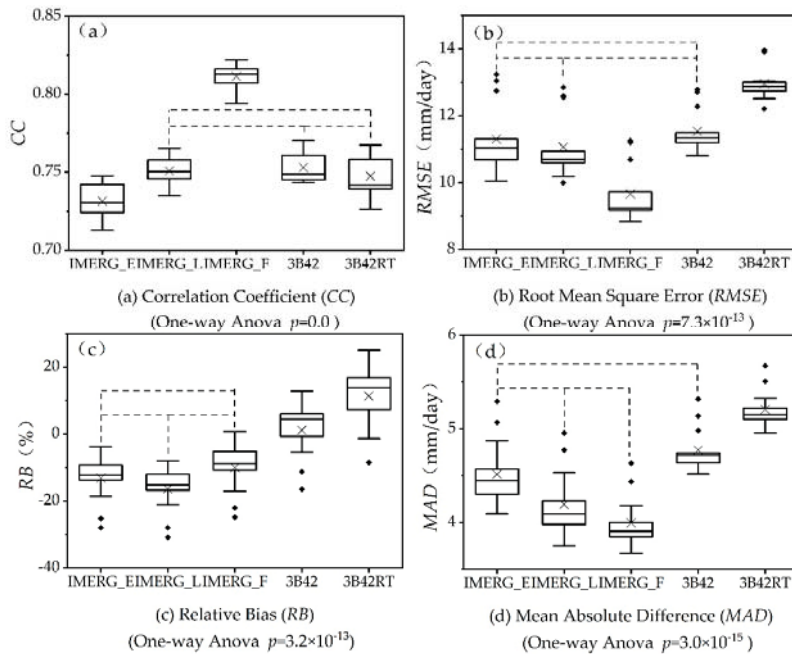
Figure 4. Boxplots of the four annual continuous evaluation metrics of SPPs at daily scale: (**a**) Correlation coefficient (*CC*) (Levene's Test, $p = 0.08$; One-Way ANOVA, $p = 0.0$); (**b**) Root-mean-square error (*RMSE*) (Levene's Test, $p = 0.16$; One-Way ANOVA, $p = 7.3 \times 10^{-13}$); (**c**) Relative bias (*RB*) (Levene's Test, $p = 0.78$; One-Way ANOVA, $p = 3.2 \times 10^{-13}$); and (**d**) Mean absolute difference (*MAD*) (Levene's Test, $p = 0.35$; One-Way ANOVA, $p = 3.0 \times 10^{-15}$). Two SPPs are connected with a black dotted line if posterior comparison tests have concluded a non-significant difference between their means at $\alpha = 0.05$. Each boxplot is used to depict the distribution, therefore the variation, of the continuous evaluation metrics among the 13 rainfall stations. In each boxplot, the top and bottom of the box represent the first and third quantiles. The whiskers extends to 1.5 times of the inter-quantile range. The horizontal line inside the box represents the median. The '×' inside the box represents the mean.

(3) Spatial Variation

Figure 5 compares the spatial distribution of the four annual continuous evaluation metrics at the daily scale among the SPPs. The spatial distribution of the CCs varies considerably among the IMERG products, while staying similar between the TMPA products. Topography does not seem to be a significant influencing factor of the CCs, although all five SPPs have lower CCs at some stations of higher altitude (station 12 and 13). In terms of RMSEs and MADs, however, topography plays a more eminent role. For all five SPPs, both metrics tend to get larger at higher altitude. In particular, RMSEs and MADs of three stations at high altitude (stations 10, 12, and 13) consistently surpass those of the rest stations. In addition, topography seems to affect the RBs of the IMERG products considerably, which tend to underestimate daily rainfall more seriously at high altitude. The impact of topography on the RBs of the TMPA products, however, is rather mixed. The absolute RB of the 3B42RT product is actually smaller at higher altitude. Similar to our study, Wang et al. (2019) [47] also observes more serious underestimation of daily rainfall by the IMERG products at high altitude in the Hexi region deep in the hinterland of the Eurasian continent. The underestimation of rainfall at high altitude by the SPPs could owe to local precipitation augmentation induced by topographical lift.
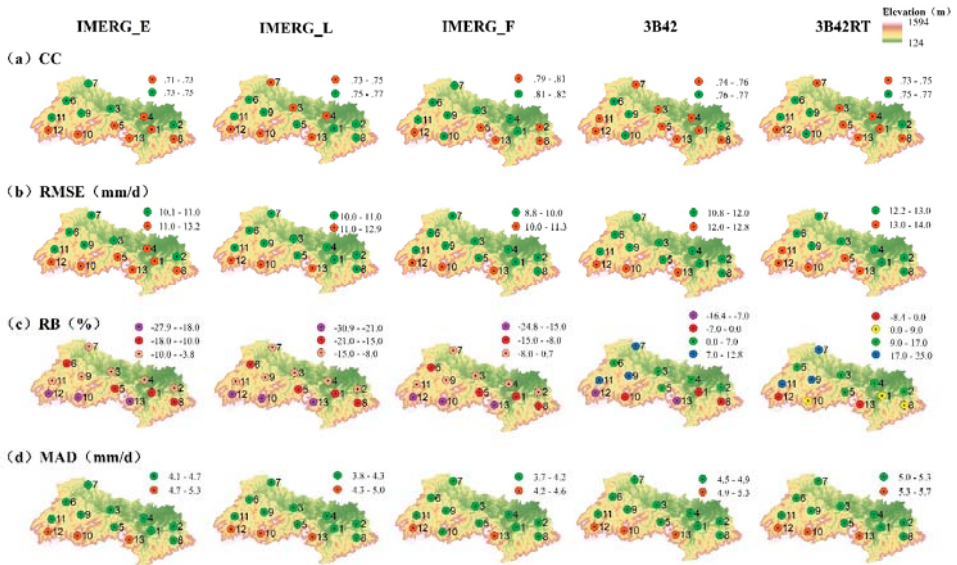
Figure 5. Spatial distribution of the annual continuous evaluation metrics at daily scale: (**a**) *CC*; (**b**) *RMSE*; (**c**) *RB*; and (**d**) *MAD*.

4.2.2. Categorical Evaluation Metrics

(1) Temporal Variation

Besides the continuous metrics, three categorical metrics are used to assess the daily precipitation detection capabilities of the SPPs. Table 4 compares the mean values of the categorical evaluation metrics among the SPPs at the daily scale. A daily rainfall threshold of 1 mm/d is used in calculating the metrics.

Table 4. Mean categorical evaluation metrics of the SPPs at daily scale.

Metrics	Temporal Scale	IMERG_E	IMERG_L	IMERG_F	3B42	3B42RT
POD	Annual	0.76	0.75	0.78	0.70	0.70
	Spring [a]	0.87	0.86	0.85	0.76	0.75
	Summer [a]	0.82	0.79	0.83	0.82	0.82
	Fall [a]	0.65	0.65	0.72	0.60	0.60
	Winter [a]	0.65	0.66	0.67	0.44	0.41
FAR	Annual	0.27	0.22	0.23	0.21	0.23
	Spring	0.23	0.17	0.19	0.12	0.12
	Summer	0.32	0.27	0.29	0.24	0.24
	Fall	0.31	0.27	0.28	0.21	0.21
	Winter	0.24	0.19	0.17	0.09	0.12
CSI	Annual	0.59	0.62	0.63	0.59	0.58
	Spring	0.69	0.73	0.71	0.69	0.68
	Summer	0.59	0.61	0.62	0.65	0.65
	Fall	0.50	0.52	0.56	0.52	0.52
	Winter	0.54	0.57	0.59	0.42	0.39

[a] Spring extends from March to May; Summer extends from June to August; Fall extends from September to November; Winter extends from December to the following February.

In terms of POD, the IMERG products all tend to perform better than the TMPA products both annually and seasonally except in summer, during which IMERG_L gives the poorest performance. Seasonally, the PODs of the two family products exhibit a somewhat different pattern of change. The PODs of the IMERG family products tend to peak (≥ 0.85) in spring, decline slightly to around 0.8 in summer, and further to around 0.67 in fall and winter. The PODs of the TMPA products peak in summer, drop to around 0.75 in spring, and further down to 0.6 in fall and < 0.45 in winter. IMERG_F has the highest POD throughout the year, except it is slightly less than the other two IMERG products in spring (Table 4). The higher PODs in spring and summer indicate that the SPPs are poorer at detecting light precipitation that is more dominant in fall and winter. Meanwhile, the much lower PODs of the TMPA products in winter indicate that it is less capable of estimating solid precipitation than the IMERG products.

Unlike POD, the IMERG products all tend to perform worse than the TMPA products in terms of FAR. Meanwhile, the FARs of the five SPPs show a similar seasonal pattern of change, which peak in summer, decrease slightly in fall, and drop further in spring and winter (Table 4).

Incorporating both correct rainfall detection and false alarm, CSI indicates a mixed performance among the SPPs. Among the five SPPs, IMERG_F performs the best annually, as well as in fall and winter. It performs slightly worse than both TMPA products in summer and IMERG_L in spring. Seasonally, all five SPPs tend to perform the best in spring and then in summer. However, the IMERG products tend to perform slightly better in winter than in fall, while the TMPA products perform considerably worse in winter (Table 4).

Similar to our study, Xu et al. (2019) [48] concludes that IMERG_F performs better than 3B42 in detecting precipitation events in the relatively flat Huang-Huai-Hai Plain of East Coastal China, with an annual POD of 0.83 and CSI of 0.52. The PODs and CSIs of IMERG_F surpass those of 3B42 in all seasons, especially in winter. This indicates that IMERG_F performs better in detecting precipitation events, especially in capturing light or solid precipitation.

(2) Spatial Variation

Figure A5 compares the spatial distribution of the three annual categorical evaluation metrics among the SPPs. Unlike the case of continuous metrics, topography does not seem to impose a consistent impact on categorical metrics at the daily scale. For example, while most SPPs have lower correct precipitation detection rates (PODs) at the two stations of high altitude (Station 12 and 13), they also have lower false alarm rates (FARs) at the stations. The leads to no obvious pattern in the spatial distribution of CSIs, with varied performance of stations at similar altitudes.
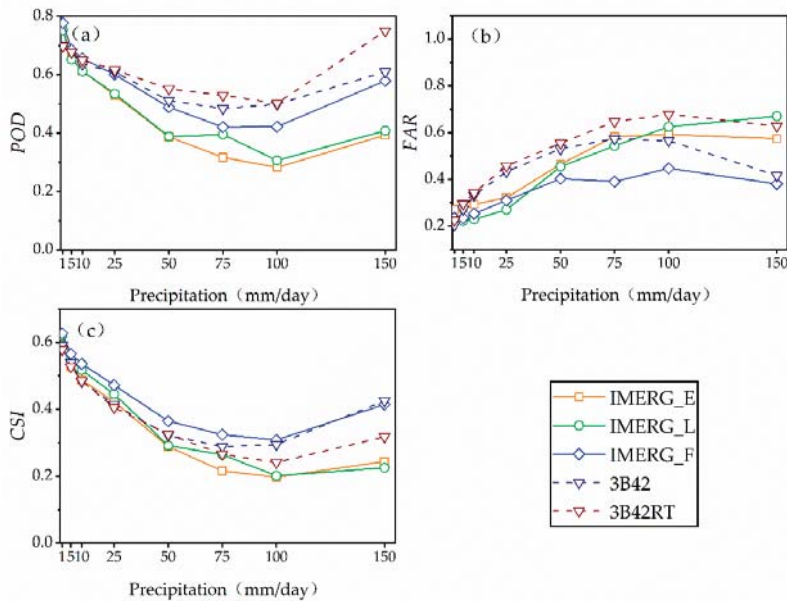
(3) Variation with Rainfall Thresholds

Figure 6 compares the performance of precipitation detection among the five SPPs by different daily rainfall magnitude. Each of the three categorical metrics has been sequentially calculated annually for the days when daily rainfall exceeds 1, 5, 10, 25, 50, 75, 100, and 150 mm/d. Similar daily rainfall thresholds have been used in previous studies, such as Wu et al. [29], Anjum et al. [43], and Tan et al. [49].

As seen from Figure 6a, the PODs of all five SPPs exhibit a largely decreasing trend with the increase of daily rainfall threshold until hitting the bottom at the threshold of 100 mm/d. Afterwards, the PODs of all SPPs bounce back substantially at the threshold of 150 mm/d. Interestingly, the PODs of the two TMPA products have mostly surpassed those of the IMERG products, indicating their better capabilities of correctly detecting daily rainfall occurrences.

However, as shown in Figure 6b, the FAR values of the TMPA products have also surpassed those of the IMERG products, especially IMERG_F, at the majority of daily rainfall thresholds, indicating their higher risk of falsely detecting daily rainfall occurrences. By incorporating the factors of both false alarms and missed events, CSI provides a more comprehensive evaluation of precipitation rainfall detection performance of the SPPs. As shown in Figure 6c, the IMERG_F has the highest CSI value at

the daily rainfall thresholds of less than 100 mm/d, whereas it is caught up by the 3B42 at the threshold of 100 mm/d and above.

Figure 6. Comparison of the changes in annual categorical evaluation metrics with daily rainfall thresholds among five SPPs: (**a**) *POD*; (**b**) *FAR*; and (**c**) *CSI*.

4.2.3. Comparison with Previous Studies

Table 5 summarizes the performance of the SPPs in estimating daily rainfall in previous studies worldwide. Previous studies have mostly assessed SPPs over approximately two years, compared to nine years in this study. It needs to be noted that the table does not serve to rigorously compare the relative performance of the SPPs in various regions, due to the differences in temporal frame, geographical regions, as well as climatic regimes of the studies.

The CCs of both IMERG products and TMPA products in this study have surpassed those in all previous studies except the one by Su et al. [46] conducted in the Upper Huai River Basin of China. Unlike the CC, the values of the other continuous as well as categorical metrics in this study all lie at the medium level among the previous studies. In addition, similar to our findings, many of previous studies have concluded a moderately better performance of the IMERG products in estimating daily rainfall than the TMPA products. However, the observed tendency of under-estimating daily rainfall by the IMERG products and over-estimating by the TMPA products in this study is not consistent with the findings of some previous studies.

Table 5. Comparison of the evaluation metrics of the SPPs at daily scale between this study and previous studies.

Study	Region	Study Area	Period	IMERG Products							TMPA Products						
				CC	RMSE (mm/d)	RB(%)	MAD (mm/d)	POD	FAR	CSI	CC	RMSE (mm/d)	RB (%)	MAD (mm/d)	POD	FAR	CSI
This study	Shuaishui River Basin, China	1.5×10^3 km^2	2009.1–2017.12	0.73–0.81	9.66–11.07	−16.33--9.99	4.00–4.52	0.75–0.78	0.22–0.27	0.59–0.63	0.75	11.54–12.96	1.21–11.36	4.77–5.21	0.70	0.21–0.23	0.58–0.59
Anjum et al. [43]	Northern Highlands, Pakistan	NA	2014.4–2016.12	0.63	1.77	7.04	/	0.72	0.24	0.54	0.70	1.85	14.77	/	0.66	0.28	0.52
Aslami et al. [50]	Ardabil Province, Iran	1.8×10^6 km^2	2016.1–2017.10	0.33	0.83	24.46	0.10	0.91	0.16	0.78	/	/	/	/	/	/	/
Beaufort et al. [31]	French Guiana	8.6×10^4 km^2	2015.4–2016.3	/	13.80	−3.10	7.40	0.70	0.36	/	/	14.40	−3.30	8.40	0.30	0.35	/
Kim et al. [51]	Japan	3.8×10^5 km^2	2014.3–2014.8	/	18.77	/	6.16	0.68	0.20	0.58	/	20.29	/	6.70	0.62	0.29	0.49
Sharifi et al. [20]	Several Provinces, Iran	6.4×10^4 km^2	2014.3–2015.2	0.40–0.52	6.38–19.41	−0.37–0.10	4.42–11.59	0.46–0.70	0.43–0.59	0.29–0.42	0.27–0.47	7.64–19.59	−9.75--1.47	5.39–11.68	0.39–0.56	0.55–0.71	0.23–0.33
Su et al. [46]	Huaihe River Basin, China	1.6×10^4 km^2	2014.4–2015.12	0.82–0.87	4.31–5.15	7.99–9.41	/	0.82–0.83	0.31–0.38	0.55–0.60	/	/	/	/	/	/	/
Tan et al. [19]	Singapore	720 km^2	2014.4–2016.6	0.53	11.83	5.24	/	0.78	0.28	0.60	0.56	9.20	−10.25	/	0.66	0.15	0.65
Tan et al. [49]	Malaysia	3.3×10^4 km^2	2014.3–2016.2	0.50–0.60	12.94–14.93	13.24	/	0.86–0.89	0.18–0.20	0.73–0.74	0.57	13.60	9.98	/	0.85	0.15	0.74
Wang et al. [52]	Mekong River Basin	7.95×10^5 km^2	2014.4–2016.1	0.58	2.24	−0.084	/	0.73	0.22	0.61	0.44	2.52	−0.10	/	0.65	0.25	0.53
Wu et al. [29]	Yangtze River Basin, China	1.8×10^6 km^2	2014.4–2017.12	0.44	11.42	5.39	4.43	/	/	/	0.46	10.75	2.87	4.10	/	/	/

4.3. Evaluation at the Hourly Scale

Figure A6 compares the scatterplots between observed and estimated hourly rainfall among the SPPs. Hourly rainfall estimates by all five SPPs are much scattered around the 1:1 line. Since hourly rainfall is only recorded from April to October, we evaluate the performance of the SPPs at the hourly scale for these seven months. Correspondingly, seasonal evaluation metrics are only calculated for spring (April to May), summer (June to August), and fall (September to October).

4.3.1. Continuous Evaluation Metrics

(1) Temporal Variation

Table 6 compares the mean values of the four continuous (CC, RMSE, RB, and MAD) over the seven months from April to October and seasonally (spring, summer, fall) among the SPPs at the hourly scale. With their seven-month CC values all staying close to 0.5, SPPs have performed less satisfactorily in estimating hourly rainfall in the SRB.

Table 6. Mean continuous evaluation metrics of the SPPs at hourly scale.

| Metrics	Temporal Scale	IMERG_E	IMERG_L	IMERG_F	3B42	3B42RT
CC	Apr. to Oct.	0.47	0.51	0.51	0.48	0.48
	Spring [a]	0.53	0.60	0.58	0.52	0.51
	Summer [a]	0.46	0.49	0.49	0.47	0.48
	Fall [a]	0.35	0.38	0.41	0.40	0.36
RMSE (mm)	Apr. to Oct.	1.57	1.51	1.56	1.60	1.65
	Spring	1.47	1.36	1.41	1.59	1.58
	Summer	1.94	1.89	1.97	1.94	2.03
	Fall	0.89	0.88	0.86	0.86	0.90
RB (%)	Apr. to Oct.	−22.18	−25.06	−10.31	−1.85	10.16
	Spring	−21.33	−24.61	−12.27	0.11	1.42
	Summer	−22.98	−26.08	−9.26	−2.20	18.35
	Fall	−20.19	−20.94	−9.77	−5.06	−7.17
MAD(mm)	Apr. to Oct.	0.33	0.31	0.33	0.35	0.37
	Spring	0.33	0.30	0.33	0.36	0.37
	Summer	0.44	0.42	0.45	0.48	0.52
	Fall	0.15	0.14	0.14	0.15	0.15

[a] Spring extends from April to May; Summer extends from June to August; Fall extends from September to October.

All IMERG products tend to underestimate hourly rainfall throughout the three seasons. The absolute RBs of IMERG_F are all below 13%, while those of the near-real-time products remain above 20%. With the smallest absolute RB among the SPPs, 3B42 tends to slightly overestimate daily rainfall in spring, but underestimate in summer and fall. Different from the other SPPs, 3B42RT shows a strong tendency of overestimation in summer.

Except for RB, the five SPPs all exhibit similar season patterns of change. Seasonally, the CCs of the SPPs all peak in spring followed by a continuous decline in summer and fall. Both the RMSEs and MADs of the SPPs are the highest in summer, followed by spring and then fall. The observed seasonal patterns at the hourly scale are quite similar to those observed at the daily scale.

Figure 7a to Figure 7l further examine the changes in mean continuous metrics over a diurnal cycle in three seasons. Meanwhile, Figure 7m to Figure 7o compare the observed amount of average hourly rainfall with the corresponding SPP estimates in the three seasons. The CCs of all five SPPs have shown considerable diurnal variations in the three seasons. Despite the differences in amount, the overall diurnal patterns of change in CCs are somewhat similar among the SPPs. In summer, for example, the CCs of all SPPs tend to reach a high plateau between 3:00 a.m. and 12:00 p.m., followed by a steady fall to the bottom at around 4:00 p.m. and a rebound afterwards. As shown in Figure 7n,

mean hourly summer ground measurement peaks at 3:00 p.m. All SPPs, however, have exhibited a lag of one or more hours in reaching the peak value, which may have caused their CCs all drop to the lowest in the afternoon. The diurnal patterns of change in RMSE/MAD are even more similar among the SPPs in all three seasons. Diurnal variations in both metrics are the highest in summer, followed by spring and then fall, which are consistent with the three seasons' relative magnitude of diurnal changes in hourly precipitation (Figure 7m–o). In addition, the RMSEs and MADs of all five SPPs peak at 5:00 a.m., 5:00 p.m., and 10:00–11:00 p.m. in spring, and at 3:00–4:00 p.m. in summer. As seen from Figure 7m,n, hourly ground measurements also peak at these times.
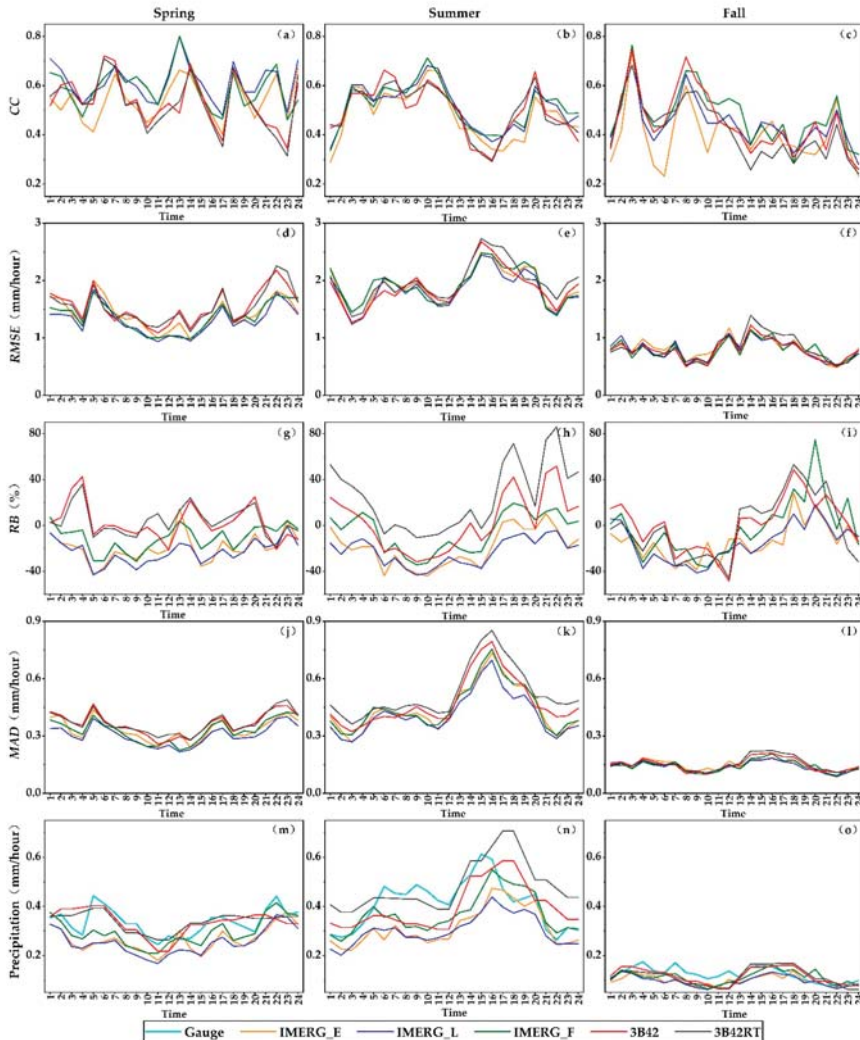
Figure 7. Changes in mean continuous metrics over a diurnal cycle in three seasons: (**a**) *CC* in spring; (**b**) *CC* in summer; (**c**) *CC* in fall; (**d**) *RMSE* in spring; (**e**) *RMSE* in summer; (**f**) *RMSE* in fall; (**g**) *RB* in spring; (**h**) *RB* in summer; (**i**) *RB* in fall; (**j**) *MAD* in spring; (**k**) *MAD* in summer; (**l**) *MAD* in fall; (**m**) average observed and estimated hourly rainfall in spring; (**n**) average observed and estimated hourly rainfall in summer; and (**o**) average observed and estimated hourly rainfall in fall.

The diurnal patterns of change in RB are more complex. Although differing much in their actual amount, the RBs of the five SPPs seem to follow a somewhat similar trend of change throughout the diurnal cycle, especially in summer. This is probably because rainfall estimates by the SPPs all exhibit a largely similar hourly trend in each season, in spite of the differences in their actual amount. Nevertheless, precisely because of the difference in their actual RB amount, the five SPPs give quite different estimation performance across the diurnal cycle. For example, 3B42RT tend to overestimate hourly rainfall most seriously at night (6:00 p.m. and 9:00–10:00 p.m.), while giving the estimates with the least bias in the morning. In contrast, IMERG products tend to underestimate rainfall mostly seriously in the morning (8:00–11:00 a.m.), but give the estimates with the least bias at night (7:00–10:00 p.m.) (Figure 7).

In addition, the relative performance of the SPPs at the hourly scale is somewhat different from that at the monthly and daily scales. In general, there is much less variability in the performance of the SPPs at the hourly scale compared to that at the monthly and daily scales. Except for RB, only two IMERG products (IMERG_F and IMERG_L) have slightly outperformed the TMPA products for most of the time.

To date, only limited studies have evaluated the quality of the hourly rainfall estimates of the SPPs. Similar to our study, they have mostly found that the performance of SPPs in estimating hourly rainfall was less satisfactory. For example, Caracciolo et al. [53] calculated the CCs to be respectively 0.32 and 0.26 when using the IMERG_F V4 for estimating hourly rainfall in Sardinia and Sicily of Italy. Li et al. [54] evaluated the performance of IMERG_F in estimating hourly rainfall in the Ganjiang River Basin of China, and calculated its CC, RMSE, and RB to be 0.33, 1.72 mm/h, and 0.12%, respectively. Yuan et al. [55] evaluated the 3-hour rainfall estimates by the three IMERG and two TMPA products in the Chindwin River basin, Myanmar, and they found that IMERG_F performed best with a CC of 0.33 and RB of −6.8%. Meanwhile, the RMSEs of the SPPs were similar, ranging from 2.9 to 3.1 mm/h.

(2) Statistical Performance Comparison among the SPPs

Figure 8 compares the boxplots of the four continuous evaluation metrics (April to October) at the hourly scale among the SPPs. For example, Figure 8b contains five boxplots, which respectively characterize the distribution of RMSEs (April to October) of the five SPPs across the 13 rainfall stations. The RMSEs of the IMERG_E range from 1.42 to 1.77 among the 13 rainfall stations, compared to 1.34–1.71 for IMERG_L, 1.45–1.74 for IMERG_F, 1.49–1.77 for 3B42, and 1.55–1.83 for 3B42RT.

Levene's Test has confirmed that all four metrics could meet the pre-condition of homogeneity of variance for conducting one-way ANOVA. The subsequent one-way ANOVA has concluded that the mean values of all four metrics are significantly different among the SPPs at the significance level (α) of 0.05. Further posterior comparison tests have shown that the CCs are significantly different between most of the pairs of SPPs except four pairs (IMERG_L/IMERG_F; IMERG_E/3B42; IMERG_E/3B42RT; and 3B42/3B42RT), while the RBs are significantly different except between two pairs (IMERG_E/IMERG_L and IMERG_F/3B42). Unlike CC and RB, the RMSEs of the SPPs are only significantly different between one pair (IMERG_L/3B42RT). Finally, the MADs are only significantly different between 3B42RT and all IMERG products, as well as between 3B42 and IMERG_L (Figure 8). It is worth noting that the posterior comparison tests have shown that IMERG_F is not significantly different from the TMPA products in terms of all four metrics except CC.
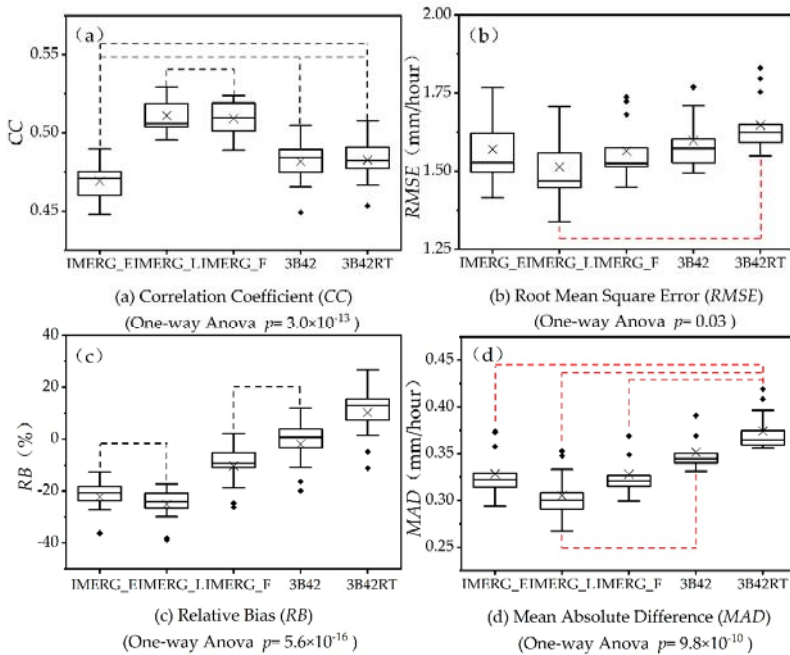
Figure 8. Boxplots of the four continuous evaluation metrics (April to October) of SPPs at hourly scale: (**a**) correlation coefficient (CC) (Levene's test, $p = 0.77$; one-way ANOVA, $p = 3.0 \times 10^{-13}$); (**b**) root-mean-square error (RMSE) (Levene's test, $p = 0.63$; one-way ANOVA, $p = 0.03$); (**c**) relative bias (RB) (Levene's test, $p = 0.61$; one-way ANOVA, $p = 5.6 \times 10^{-16}$); and (**d**) mean absolute difference (MAD) (Levene's test, $p = 0.87$; one-way ANOVA, $p = 9.8 \times 10^{-10}$). For CC and RB, two SPPs are connected with a black dotted line if posterior comparison tests have indicated a non-significant difference between their means at $\alpha = 0.05$. For RMSE and MAD, two SPPs are connected with a red dotted line if posterior comparison tests have indicated a significant difference between their means at $\alpha = 0.05$. Each boxplot is used to depict the distribution, therefore the variation, of the continuous evaluation metrics among the 13 rainfall stations. In each boxplot, the top and bottom of the box represent the first and third quantiles. The whiskers extends to 1.5 times of the inter-quantile range. The horizontal line inside the box represents the median. The '×' inside the box represents the mean.

(3) Spatial Variation

Figure 9 compares the spatial distribution of the four annual continuous evaluation metrics among the SPPs. At the hourly scale, topography also does not seem to be a significant influencing factor of the CCs, with lower CC values observed at stations of both low and high altitude. However, the spatial distribution of the other three metrics does indicate a significant impact of topography on the performance of the SPPs in estimating hourly rainfall. Both RMSEs and MADs exhibit similar spatial patterns across the five SPPs, whose values at the three stations of high altitude (station 10, 12, and 13) consistently stay at the top. As discussed above, the IMERG products tend to underestimate hourly rainfall. As shown in Figure 9, underestimation by the IMERG products is especially severe at higher altitude. Meanwhile, the 3B42 product also tends to underestimate hourly rainfall more at high altitude. Different from the other SPPs, the 3B42RT product tends to overestimate hourly rainfall more seriously at lower altitude.
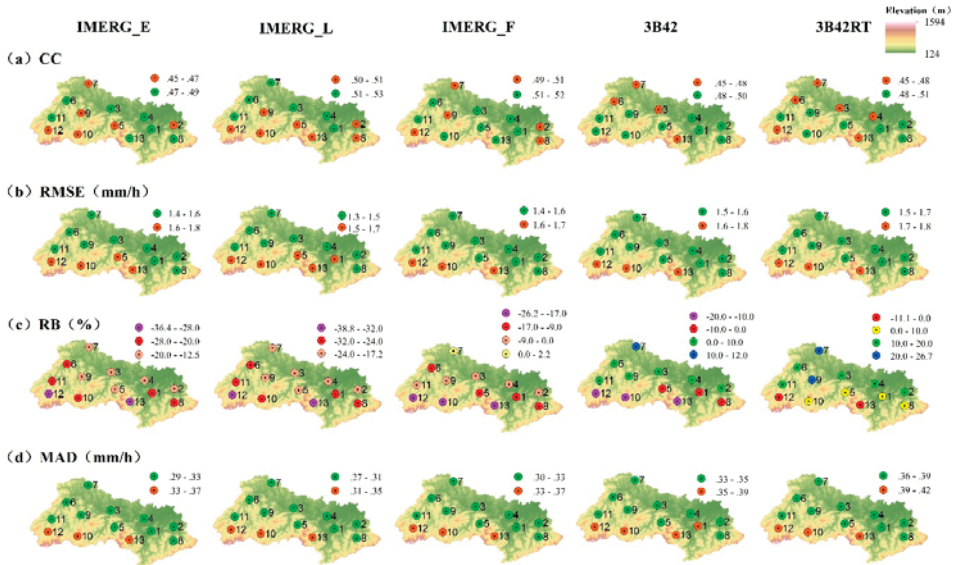
Figure 9. Spatial distribution of the continuous evaluation metrics (April to October) at hourly scale: (**a**) *CC*; (**b**) *RMSE*; (**c**) *RB*; and (**d**) *MAD*.

4.3.2. Categorical Evaluation Metrics

(1) Temporal Variation

Table 7 compares the mean values of the three categorical metrics (POD, FAR, and CSI) over the seven months from April to October and seasonally (spring, summer, fall) among the SPPs at the hourly scale. Similar to the daily scale, three categorical metrics are used to assess the hourly precipitation detection capabilities of the SPPs. An hourly rainfall threshold of 0.1 mm/d is used in calculating the metrics. As seen from the table, with lower PODs and CSIs, as well as higher FARs, all five SPPs are poorer at detecting hourly rainfall than daily rainfall.

**Table 7.** Mean categorical evaluation metrics of the SPPs at hourly scale.

Metrics	Temporal Scale	IMERG_E	IMERG_L	IMERG_F	3B42	3B42RT
POD	Apr. to Oct.	0.63	0.66	0.67	0.65	0.65
	Spring [a]	0.65	0.68	0.65	0.63	0.63
	Summer [a]	0.66	0.69	0.71	0.71	0.71
	Fall [a]	0.49	0.54	0.60	0.53	0.54
FAR	Apr. to Oct.	0.50	0.49	0.50	0.50	0.51
	Spring	0.46	0.45	0.43	0.40	0.41
	Summer	0.53	0.50	0.52	0.54	0.55
	Fall	0.51	0.49	0.53	0.50	0.51
CSI	Apr. to Oct.	0.38	0.41	0.40	0.39	0.39
	Spring	0.42	0.44	0.43	0.44	0.43
	Summer	0.38	0.41	0.40	0.38	0.38
	Fall	0.32	0.35	0.36	0.35	0.35

[a] Spring extends from April to May; Summer extends from June to August; Fall extends from September to October.

Seasonally, the five SPPs have exhibited somewhat similar patterns of change in hourly rainfall detection performance. In summer, they all have the highest correct rainfall detection rates (PODs),

but also the highest false alarm rates (FARs). Between the rest two seasons, all five SPPs have higher PODs as well as lower FARs, therefore better rainfall detection performance, in spring. In fact, the seasonal CSIs indicate that the overall rainfall detection performance of the SPPs all tops in spring, followed by summer, and then winter.

Figure 10 further examines the changes in mean categorical metrics over a diurnal cycle in three seasons. Among the three categorical metrics, PODs, especially those of IMERG_L and IMERG_F, exhibit relatively less hourly variations through the diurnal cycle. The only discernible pattern in the metric is that the IMERG products tend to have the highest correct rainfall detection rates (> 0.7) in early evening, while the TMPA products have the lowest at midnight in spring. Unlike POD, FAR exhibits more diurnal variations. In spring, FARs of all five SPPs tend to peak around noon. Whereas, in summer and fall, they all tend to bottom in the morning and climb to the peak at around midnight. Similar to FAR, CSI exhibits distinct diurnal variations. In spring, all SPPs have the lowest CSI, i.e., the poorest hourly rainfall detection performance at noon and midnight. In summer and fall, the performance of the SPPs tends to peak in the morning, and bottom out around midnight.
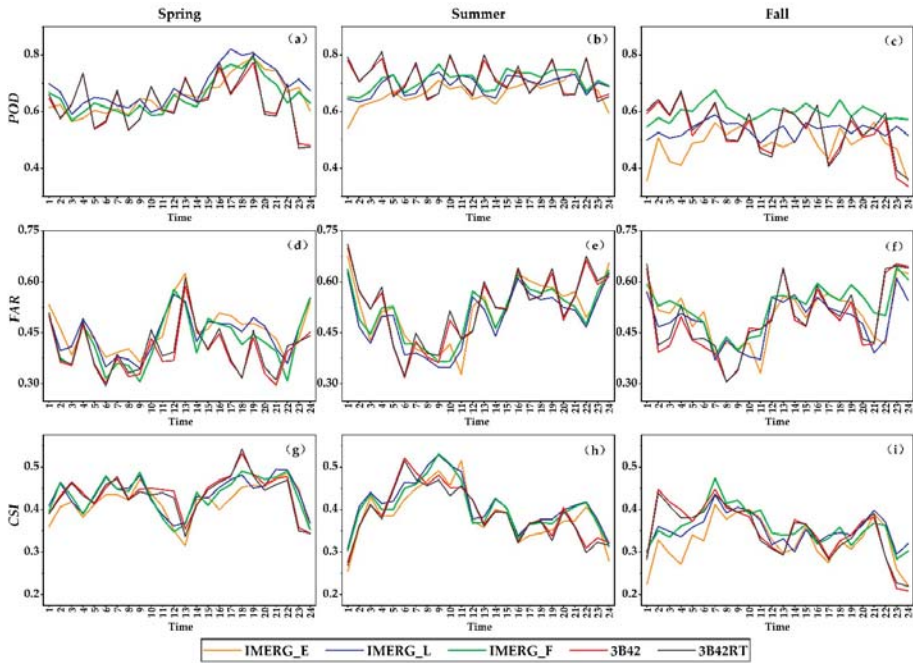
Figure 10. Changes in mean categorical metrics over a diurnal cycle in three seasons: (**a**) *POD* in spring; (**b**) *POD* in summer; (**c**) *POD* in fall; (**d**) *FAR* in spring; (**e**) *FAR* in summer; (**f**) *FAR* in fall; (**g**) *CSI* in spring; (**h**) *CSI* in summer; and (**i**) *CSI* in fall.

Similar to the results of the continuous metrics at the hourly scale, the values of all three categorical metrics are not much different among the SPPs, indicating the IMERG products are no longer superior to the TMPA products in hourly rainfall detection.

The values of the categorical metrics obtained in this study lie at the medium level among those obtained in previous limited studies. For example, Omranian and Sharif [56] evaluated the Early, Late, and Final runs of IMERG products for May 2015 in the Lower Colorado River Basin of Texas in U.S. At the hourly scale, they obtained a POD value of 0.74, a FAR value of 0.45 and a CSI value of 0.46. However, in their evaluation of the 3-hour daytime satellite rainfall estimates in the Chindwin

River basin, Myanmar, Yuan et al. [55] obtained a POD value of <0.3 for all three IMERG and two TMPA products.

(2) Spatial Variation

Figure A7 compares the spatial distribution of the three annual categorical evaluation metrics among the SPPs. At the hourly scale, the spatial distribution of the categorical metrics exhibits a similar pattern across the five SPPs. All five SPPs are poorer at detecting the actual precipitation occurrences at high altitude. However, they also tend to yield more false alarms at low altitude. The conflicting impacts of elevation on the PODs and FARs of the SPPs have led to lower CSI values, i.e., worse overall hourly precipitation detection performance, at the lower altitude, especially at stations 1–3.

(3) Variation with Rainfall Thresholds

Figure 11 compares the performance of hourly precipitation detection among the five SPPs by different rainfall magnitude. Each of the three categorical metrics has been sequentially calculated when hourly rainfall exceeds 0.1, 1, 5, 10, and 15 mm. As seem from Figure 11a, the PODs of the five SPPs have all decreased steadily with increasing hourly rainfall thresholds before plunging to nearly 0 at the threshold of 15 mm/hour. In general, the five SPPs do not differ much in their PODs across the entire range of hourly rainfall thresholds. The PODs of the two near-real-time IMERG products are consistently less than those of the TMPA products. The PODs of 3B42RT have actually remained at or nearly the top across the rainfall thresholds.
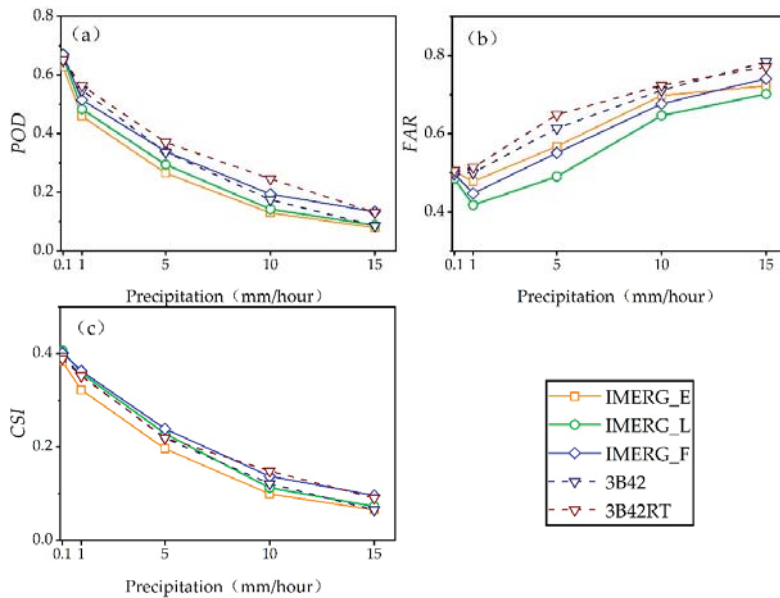

Figure 11. Comparison of the changes in categorical evaluation metrics (April to October) with hourly rainfall thresholds among five SPPs: (**a**) *POD*; (**b**) *FAR*; and (**c**) *CSI*.

As seem from Figure 11b, the FARs of all five SPPs tend to rise with increasing rainfall thresholds. The TMPA products have consistently yielded higher FARs than the IMERG products over the entire range of rainfall thresholds. The more comprehensive CSI values have concluded much similar hourly rainfall detection performance among the SPPs across the rainfall thresholds, with IMERG_F staying at or near the top most of the time (Figure 11c).

5. Conclusions

SPPs have increasingly become an important data source for precipitation inputs in hydrological modeling and other related studies worldwide. For local regions with scarce precipitation observations or limited access to precipitation data, the latest GPM and TRMM products provide a valuable alternative for obtaining the much-needed rainfall inputs for various regional hydrological applications. However, the accuracy of their rainfall estimates should be systematically assessed before being utilized in real world applications. In this study, we have assessed and compared the accuracy of the latest five GPM IMERG V6 and TRMM 3B42 V7 precipitation products across the monthly, daily, and hourly scales in a middle-sized hilly river basin in eastern central China. For evaluation, a total of four continuous and three categorical metrics have been calculated based on SPP estimates and historical rainfall records at 13 stations over a period of 9 years from 2009 to 2017. The evaluation results have led to the following main conclusions:

(1) Rainfall estimates by all five SPPs could match ground observations best at the monthly scale, followed by the daily and hourly scale. The annual CCs of the SPPs, for example, have fallen from 0.86 or above at the monthly scale to mostly around 0.75 at the daily scale, and sharply to less than 0.6 (April to October) at the hourly scale. Topography tends to impose similar impact on the performance of SPPs across various time scales, with more estimation deviations at high altitude.

(2) For estimating monthly rainfall, IMERG_F performs the best, closely followed by 3B42. These two post-time SPPs produce considerably better monthly rainfall estimates than the rest real-time or near-real-time SPPs. All three IMERG products tend to underestimate monthly rainfall except a slight overestimation by the two near-real-time products in winter. Meanwhile, 3B42RT exhibits a strong tendency to overestimate in summer and winter.

(3) For estimating daily rainfall, the IMERG products generally perform better than the TMPA products, with IMERG_F performing the best. Similar to the monthly scale, the IMERG family products tend to underestimate daily rainfall in all four seasons except the two near-real-time products in winter. In contrast, 3B42RT exhibits a strong tendency of overestimation in summer and winter. In terms of rainfall detection performance, the TMPA products are more capable of correctly detecting daily rainfall occurrences, while the IMERG products contain fewer false detections of rainfall occurrences.

(4) For estimating hourly rainfall, the performance of the SPPs is much more homogeneous. Two IMERG products (IMERG_F and IMERG_L) have slightly outperformed the TMPA products for most of the time. All IMERG products tend to underestimate hourly rainfall throughout the three seasons between April and October. In contrast, 3B42RT shows a strong tendency of overestimation in summer. In addition, the performances of hourly rainfall detection are quite similar among the five SPPs.

In general, our nine-year systematic evaluation of the latest GPM IMERG V6 and TRMM 3B42 V7 precipitation products have shown that the SPPs, especially the post-time IMERG_F product, could be considered as a reliable data source for providing monthly or daily rainfall data for regional hydrological applications. However, great caution needs to be exerted to utilize the hourly rainfall SPPs considering their overall weak correlations with ground rainfall observations, as well as the consistent tendency of underestimation by the IMERG products.

Hourly rainfall datasets have been increasingly found to be valuable inputs to a variety of hydrological applications. However, limited access to hourly rainfall datasets have restrained such applications in many regions. Owing to their wide spatial coverage and open access, SPPs have great potential to act as a useful alternative source for providing hourly rainfall data. Therefore, effective bias-correction algorithms incorporating ground rainfall observations are needed to improve the quality of hourly rainfall SPPs to safeguard the validity of their usage as ground measurement surrogates.
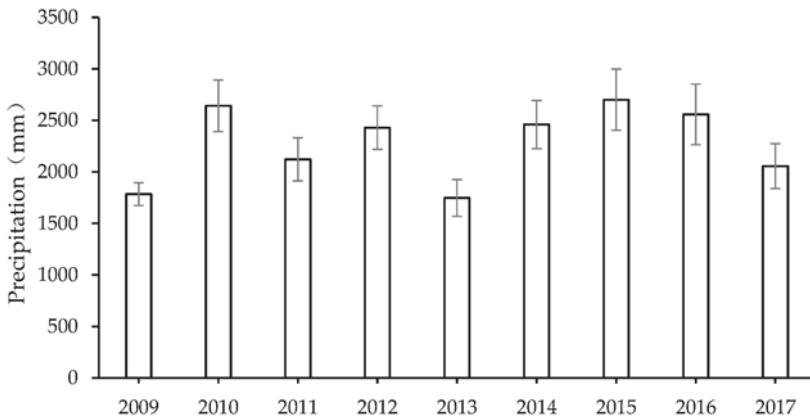
Author Contributions: Conceptualization, X.Y. and M.L.T.; Methodology, X.Y. and R.H.; Software, Y.L. and X.L.; Validation, X.Y. and G.W.; Formal analysis, X.Y. and Y.L.; Investigation, X.Y. and Y.L.; Resources, G.W. and R.H.;

Data curation, Y.L. and X.L.; Writing—original draft preparation, X.Y. and Y.L.; Writing—review and editing, M.L.T., G.W., and R.H.; Visualization, Y.L. and X.L.; Supervision, X.Y. and R.H.; Project administration, X.Y. and R.H.; Funding acquisition, R.H. and X.Y. All authors have read and agreed to the published version of the manuscript.
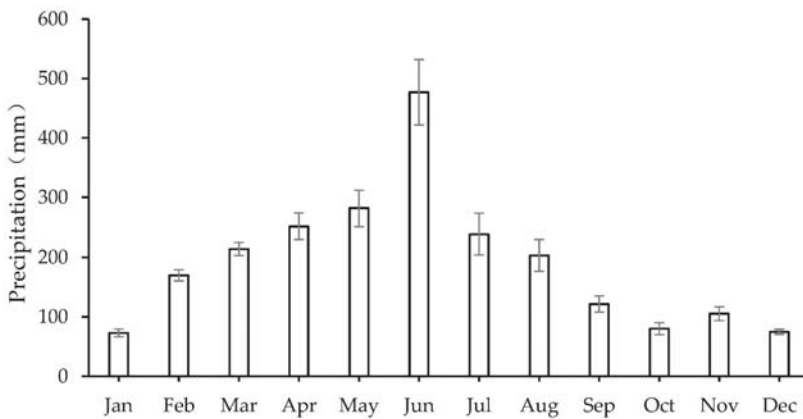
Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Figure A1. Mean annual precipitation in the Shuaishui River Basin: 2009–2017. The error bar shows ± one standard deviation of the annual average precipitation among the 13 rainfall stations.

Figure A2. Mean monthly precipitation in the Shuaishui River Basin: 2009–2017. The error bar shows ± one standard deviation of the monthly average precipitation among the 13 stations.

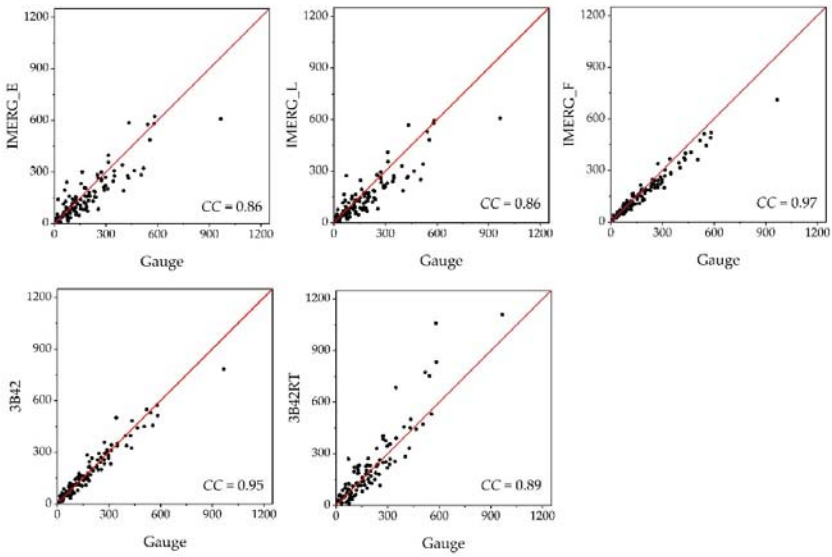**Figure A3.** Scatterplots between monthly rainfall gauge observations and satellite precipitation product estimates.
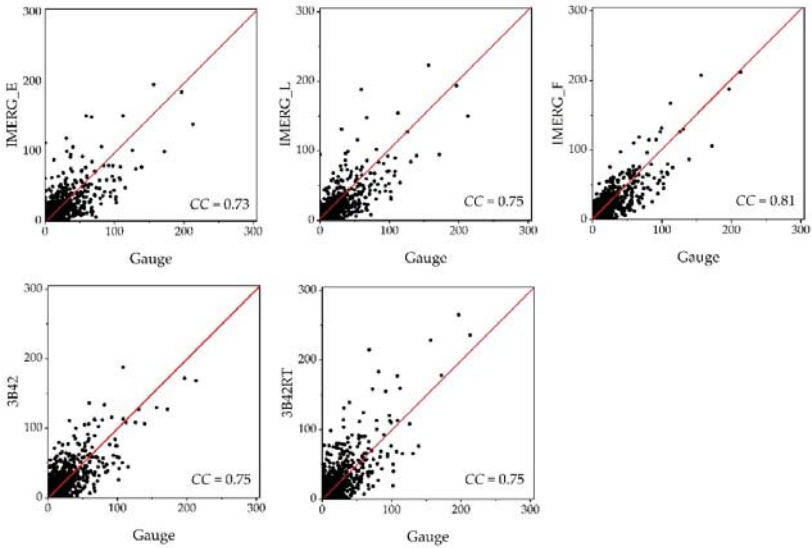

Figure A4. Scatterplots between daily rainfall gauge observations and satellite precipitation product estimates.

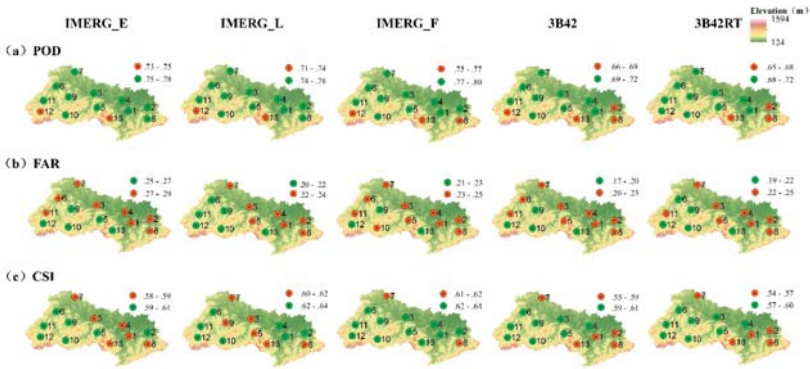**Figure A5.** Spatial distribution of the annual categorical evaluation metrics at daily scale: (**a**) *POD*; (**b**) *FAR*; and (**c**) *CSI*.
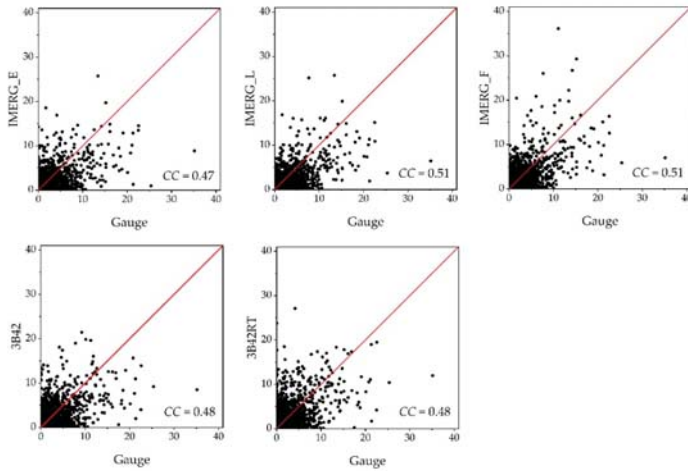

Figure A6. Scatterplots between hourly rainfall gauge observations and satellite precipitation product estimates.
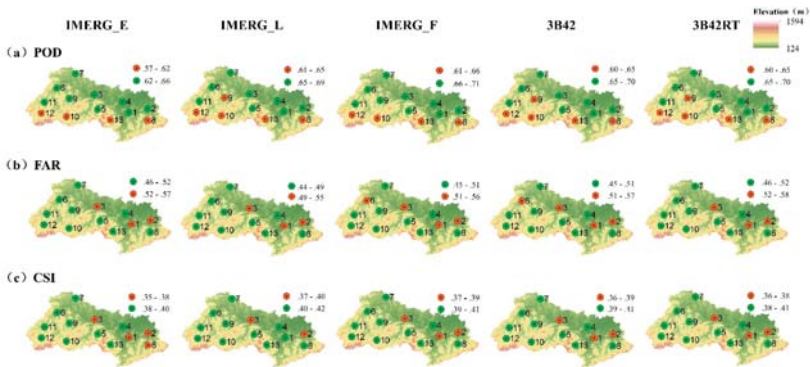

Figure A7. Spatial distribution of the categorical evaluation metrics (April to October) at hourly scale: (**a**) *POD*; (**b**) *FAR*; and (**c**) *CSI*.

References

1. Wu, P.; Christidis, N.; Stott, P. Anthropogenic impact on Earth's hydrological cycle. *Nat. Clim. Chang.* **2013**, *3*, 807–810. [CrossRef]

2. Allen, M.R.; Ingram, W.J. Constraints on future changes in climate and the hydrologic cycle. *Nature* **2012**, *489*, 590. [CrossRef]

3. Lee, J.; Lee, E.-H.; Seol, K.-H. Validation of Integrated MultisatellitE Retrievals for GPM (IMERG. by using gauge-based analysis products of daily precipitation over East Asia. *Theor. Appl. Climatol.* **2019**, *137*, 2497–2512. [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. [CrossRef]

5. Shen, Y.; Xiong, A. Validation and comparison of a new gauge-based precipitation analysis over mainland China. *Int. J. Climatol.* **2016**, *36*, 252–265. [CrossRef]

6. Tan, M.L.; Ibrahim, A.; 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]

7. Tancreto, A.E. *Comparison of Hydrologic Model Performance Statistics Using Thiessen Polygon Rain Gauge and NEXRAD Precipitation Input Methods at Different Watershed Spatial Scales and Rainfall Return Frequencies*; University of North Florida: Jacksonville, FL, USA, 2015; Available online: https://digitalcommons.unf.edu/etd/584 (accessed on 23 March 2020).

8. Louf, V.; Protat, A.; Warren, R.A.; Collis, S.M.; Wolff, D.B.; Raunyiar, S.; Jakob, C.; Petersen, W.A. An integrated approach to weather radar calibration and monitoring using ground clutter and satellite comparisons. *J. Atmos. Ocean. Technol.* **2019**, *36*, 17–39. [CrossRef]

9. Letu, H.; Nagao, T.M.; Nakajima, T.Y.; Riedi, J.; Ishimoto, H.; Baran, A.J.; Shang, H.; Sekiguchi, M.; Kikuchi, M. Ice cloud properties from Himawari-8/AHI next-generation geostationary satellite: Capability of the AHI to monitor the DC cloud generation process. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3229–3239. [CrossRef]

10. Rutledge, S.A.; Chandrasekar, V.; Fuchs, B.; George, J.; Junyent, F.; Kennedy, P.; Dolan, B. Deployment of the SEA-POL C-band Polarimetric Radar to SPURS-2. *Oceanography* **2019**, *32*, 50–57. [CrossRef]

11. Beck, H.E.; Pan, M.; Roy, T.; Weedon, G.P.; Pappenberger, F.; van Dijk, A.I.J.M.; Huffman, G.J.; Adler, R.F.; Wood, E.F. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 207–224. [CrossRef]

12. Joss, J.; Waldvogel, A.; Collier, C.G. Precipitation Measurement and Hydrology. In *Radar in Meteorology: Battan Memorial and 40th Anniversary Radar Meteorology Conference*; Atlas, D., Ed.; American Meteorological Society: Boston, MA, USA, 1990; pp. 577–606. [CrossRef]

13. Dinku, T.; Anagnostou, E.N.; Borga, M. Improving radar-based estimation of rainfall over complex terrain. *J. Appl. Meteorol.* **2002**, *41*, 1163–1178. [CrossRef]

14. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A.; et al. The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*. [CrossRef]

15. 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. *Remote Sens.* **2015**, *7*, 7181–7211. [CrossRef]

16. Mantas, V.M.; Liu, Z.; Caro, C.; Pereira, A.J.S.C. Validation of TRMM multi-satellite precipitation analysis (TMPA) products in the Peruvian Andes. *Atmos. Res.* **2015**, *163*, 132–145. [CrossRef]

17. Li, X.-H.; Zhang, Q.; Xu, C.-Y. Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. *J. Hydrol.* **2012**, *426*, 28–38. [CrossRef]

18. Bharti, V.; Singh, C. Evaluation of error in TRMM 3B42V7 precipitation estimates over the Himalayan region. *J. Geophys. Res. -Atmos.* **2015**, *120*, 12458–12473. [CrossRef]

19. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]

20. Sharifi, E.; Steinacker, R.; Saghafian, B. Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sens.* **2016**, *8*, 135. [CrossRef]

21. 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]

22. Suratman, S.; Aziz, A.A.; Tahir, N.M.; Lee, L.H. Distribution and behaviour of nitrogen compounds in the surface water of Sungai Terengganu Estuary, Southern Waters of South China Sea, Malaysia. *Sains Malays.* **2018**, *47*, 651–659. [CrossRef]

23. Zhou, Y.; Wu, T. Composite analysis of precipitation intensity and distribution characteristics of western track landfall typhoons over China under strong and weak monsoon conditions. *Atmos. Res.* **2019**, *225*, 131–143. [CrossRef]

24. Yang, X.; Warren, R.; He, Y.; Ye, J.; Li, Q.; Wang, G. Impacts of climate change on TN load and its control in a river basin with complex pollution sources. *Sci. Total Environ.* **2018**, *615*, 1155–1163. [CrossRef]

25. Yang, X.; Liu, Q.; Fu, G.; He, Y.; Luo, X.; Zheng, Z. Spatiotemporal patterns and source attribution of nitrogen load in a river basin with complex pollution sources. *Water Res.* **2016**, *94*, 187–199. [CrossRef]

26. Boithias, L.; Sauvage, S.; Lenica, A.; Roux, H.; Abbaspour, K.C.; Larnier, K.; Dartus, D.; Sanchez-Perez, J.M. Simulating flash floods at hourly time-step using the SWAT model. *Water* **2017**, *9*, 929. [CrossRef]

27. Zhou, M.M.; Deng, J.S.; Lin, Y.; Belete, M.; Wang, K.; Comber, A.; Huang, L.Y.; Gan, M.Y. Identifying the effects of land use change on sediment export: Integrating sediment source and sediment delivery in the Qiantang River Basin, China. *Sci. Total Environ.* **2019**, *686*, 38–49. [CrossRef]

28. Lin, Q.W.; Peng, X.; Liu, B.Y.; Min, F.L.; Zhang, Y.; Zhou, Q.H.; Ma, J.M.; Wu, Z.B. Aluminum distribution heterogeneity and relationship with nitrogen, phosphorus and humic acid content in the eutrophic lake sediment. *Environ. Pollut.* **2019**, *253*, 516–524. [CrossRef]

29. Wu, Y.; Zhang, Z.; Huang, Y.; Jin, Q.; Chen, X.; Chang, J. Evaluation of the GPM IMERG V5 and TRMM 3B42 V7 precipitation products in the Yangtze River Basin, China. *Water* **2019**, *11*, 1459. [CrossRef]

30. Zhao, H.; Yang, S.; Wang, Z.; Zhou, X.; Luo, Y.; Wu, L. Evaluating the suitability of TRMM satellite rainfall data for hydrological simulation using a distributed hydrological model in the Weihe River catchment in China. *J. Geogr. Sci.* **2015**, *25*, 177–195. [CrossRef]

31. Beaufort, A.; Gibier, F.; Palany, P. Assessment and correction of three satellite rainfall estimate products for improving flood prevention in French Guiana. *Int. J. Remote Sens.* **2019**, *40*, 171–196. [CrossRef]

32. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.J.; Bowman, K.P.; Hong, Y.; Stocker, E.F.; Wolff, D.B. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]

33. Real-Time TRMM Multi-Satellite Precipitation Analysis Data Set Documentation. Available online: https://pmm.nasa.gov/sites/default/files/document_files/3B4XRT_doc_V7_180426.pdf (accessed on 17 March 2020).

34. Bushair, M.T.; Kumar, P.; Gairola, R.M. Evaluation and assimilation of various satellite-derived rainfall products over India. *Int. J. Remote Sens.* **2019**, *40*, 5315–5338. [CrossRef]

35. Algorithm Theoretical Basis Document (ATBD) Version 06. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf (accessed on 17 March 2020).

36. Yong, B.; Ren, L.; Hong, Y.; Gourley, J.J.; Tian, Y.; Huffman, G.J.; Chen, X.; Wang, W.; Wen, Y. First evaluation of the climatological calibration algorithm in the real-time TMPA precipitation estimates over two basins at high and low latitudes. *Water Resour. Res.* **2013**, *49*, 2461–2472. [CrossRef]

37. Condom, T.; Rau, P.; Espinoza, J.C. Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007. *Hydrol. Process.* **2011**, *25*, 1924–1933. [CrossRef]

38. Blacutt, L.A.; Herdies, D.L.; de Goncalves, L.G.G.; Vila, D.A.; Andrade, M. Precipitation comparison for the CFSR, MERRA, TRMM3B42 and Combined Scheme datasets in Bolivia. *Atmos. Res.* **2015**, *163*, 117–131. [CrossRef]

39. El Kenawy, A.M.; Lopez-Moreno, J.I.; McCabe, M.F.; Vicente-Serrano, S.M. Evaluation of the TMPA-3B42 precipitation product using a high-density rain gauge network over complex terrain in northeastern Iberia. *Glob. Planet. Chang.* **2015**, *133*, 188–200. [CrossRef]

40. Yong, B.; Ren, L.-L.; Hong, Y.; Wang, J.-H.; Gourley, J.J.; Jiang, S.-H.; Chen, X.; Wang, W. Hydrologic evaluation of multisatellite precipitation analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. *Water Resour. Res.* **2010**, *46*. [CrossRef]

41. Gerapetritis, H.; Pelissier, J. The critical success index and warning strategy. In Proceedings of the 17th Conference on Probablity and Statistics in the Atmospheric Sciences, Seattle, WC, USA, 11–15 January 2004.

42. Chen, S.; Hong, Y.; Cao, Q.; Gourley, J.J.; Kirstetter, P.-E.; Yong, B.; Tian, Y.; Zhang, Z.; Shen, Y.; Hu, J.; et al. Similarity and difference of the two successive V6 and V7 TRMM multisatellite precipitation analysis performance over China. *J. Geophys. Res. -Atmos.* **2013**, *118*, 13060–13074. [CrossRef]

43. Anjum, M.N.; Ding, Y.; Shangguan, D.; Ahmad, I.; Ijaz, M.W.; Farid, H.U.; Yagoub, Y.E.; Zaman, M.; Adnan, M. Performance evaluation of latest integrated multi-satellite retrievals for Global Precipitation Measurement (IMERG) over the northern highlands of Pakistan. *Atmos. Res.* **2018**, *205*, 134–146. [CrossRef]

44. Yang, M.; Li, Z.; Anjum, M.N.; Gao, Y. Performance Evaluation of Version 5 (V05) of Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) over the Tianshan Mountains of China. *Water* **2019**, *11*, 1139. [CrossRef]

45. Milewski, A.; Elkadiri, R.; Durham, M. Assessment and comparison of TMPA satellite precipitation products in varying climatic and topographic regimes in Morocco. *Remote Sens.* **2015**, *7*, 5697–5717. [CrossRef]

46. Su, J.; Lu, H.; Zhu, Y.; Cui, Y.; Wang, X. Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China. *Atmos. Res.* **2019**, *225*, 17–29. [CrossRef]

47. Wang, X.; Ding, Y.; Zhao, C.; Wang, J. Similarities and improvements of GPM IMERG upon TRMM 3B42 precipitation product under complex topographic and climatic conditions over Hexi region, Northeastern Tibetan Plateau. *Atmos. Res.* **2019**, *218*, 347–363. [CrossRef]

48. Xu, F.; Guo, B.; Ye, B.; Ye, Q.; Chen, H.; Ju, X.; Guo, J.; Wang, Z. Systematical Evaluation of GPM IMERG and TRMM 3B42V7 Precipitation Products in the Huang-Huai-Hai Plain, China. *Remote Sens.* **2019**, *11*, 697. [CrossRef]

49. Tan, M.L.; Santo, H. Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia. *Atmos. Res.* **2018**, *202*, 63–76. [CrossRef]

50. Aslami, F.; Ghorbani, A.; Sobhani, B.; Esmali, A. Comprehensive comparison of daily IMERG and GSMaP satellite precipitation products in Ardabil Province, Iran. *Int. J. Remote Sens.* **2019**, *40*, 3139–3153. [CrossRef]

51. Kim, K.; Park, J.; Baik, J.; Choi, M. Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.* **2017**, *187*, 95–105. [CrossRef]

52. Wang, W.; Lu, H.; Zhao, T.; Jiang, L.; Shi, J. Evaluation and comparison of daily rainfall from latest GPM and TRMM products over the Mekong River Basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2540–2549. [CrossRef]

53. Caracciolo, D.; Francipane, A.; Viola, F.; Noto, L.V.; Deidda, R. Performances of GPM satellite precipitation over the two major Mediterranean islands. *Atmos. Res.* **2018**, *213*, 309–322. [CrossRef]

54. Li, N.; Tang, G.; Zhao, P.; Hong, Y.; Gou, Y.; Yang, K. Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis IMERG in Ganjiang River basin. *Atmos. Res.* **2017**, *183*, 212–223. [CrossRef]

55. Yuan, F.; Zhang, L.; Soe, K.M.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Applications of TRMM- and GPM-Era multiple-satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in Myanmar. *Remote Sens.* **2019**, *11*, 140. [CrossRef]

56. Omranian, E.; Sharif, H.O. Evaluation of the Global Precipitation Measurement (GPM) satellite rainfall products over the Lower Colorado River Basin, Texas. *J. Am. Water Resour. Assoc.* **2018**, *54*, 882–898. [CrossRef]

Article

The Development of a Two-Step Merging and Downscaling Method for Satellite Precipitation Products

Xinyu Lu [1], Guoqiang Tang [2,3,*], Xiuqin Wang [1], Yan Liu [1], Ming Wei [4] and Yingxin Zhang [1]

[1] Institute of Desert Meteorology, China Meteorological Administration, Urumqi 830002, China;
     luxy@idm.cn (X.L.); xiuqinwang66517330@163.com (X.W.); liuyan@idm.cn (Y.L.);
     zhangyingxin@bjmb.gov.cn (Y.Z.)
[2] Coldwater Lab, University of Saskatchewan, Canmore, AB T1W 3G1, Canada
[3] Center for Hydrology, University of Saskatchewan, Saskatoon, SK S7N 1K2, Canada
[4] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University
     of Information Science & Technology, Nanjing 210044, China; mingwei@nuist.edu.cn
* Correspondence: guoqiang.tang@usask.ca

Abstract: Low accuracy and coarse spatial resolution are the two main drawbacks of satellite precipitation products. Therefore, calibration and downscaling are necessary before these products are applied. This study proposes a two-step framework to improve the accuracy of satellite precipitation estimates. The first step is data merging based on optimum interpolation (OI), and the second step is downscaling based on geographically weighted regression (GWR); therefore, the framework is called OI-GWR. An Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) product is used to demonstrate the effectiveness of OI-GWR in the Tianshan Mountains, China. First, the original IMERG precipitation data (OIMERG) are merged with rain gauge data using the OI method to produce corrected IMERG precipitation data (CIMERG). Then, using CIMERG as the first guess and the normalized difference vegetation index (NDVI) as the auxiliary variable, GWR is utilized for spatial downscaling. The two-step OI-GWR method is compared with several traditional methods, including GWR downscaling (Ori_GWR) and spline interpolation. The cross-validation results show that (1) the OI method noticeably improves the accuracy of OIMERG, and (2) the 1-km downscaled data obtained using OI-GWR are much better than those obtained from Ori_GWR, spline interpolation, and OIMERG. The proposed OI-GWR method can contribute to the development of high-resolution and high-accuracy regional precipitation datasets. However, it should be noted that the method proposed in this study cannot be applied in regions without any meteorological stations. In addition, further efforts will be needed to achieve daily- or hourly-scale downscaling of precipitation.

Keywords: precipitation; IMERG; optimum interpolation; geographically weighted regression; downscaling; Tianshan Mountains

---

1. Introduction

Precipitation is the main component of the global water cycle and plays a critical role in Earth's energy balance [1,2]. Therefore, accurate information on the spatial and temporal distribution of precipitation is essential to improve our understanding of the Earth system and to better predict weather and climate conditions and natural disasters. Ground observations are the most direct source of precipitation data, but most stations are unevenly distributed in low-altitude zones, which makes it difficult to capture the full distribution of large-scale precipitation. In contrast, satellite-retrieved precipitation products have the unique advantages of global coverage and spatiotemporal continuity [3,4] and have

consequently promoted a more complete understanding of the patterns of and changes in regional and global precipitation.

However, because of instrument limitations and imperfect retrieval algorithms, satellite precipitation products have drawbacks in terms of spatial resolution and data precision [5–9]. At present, because of the coarse spatial resolution of satellite precipitation products, their application in hydrological and climatic models at the watershed scale is restricted. For this reason, many researchers have focused on developing statistical downscaling methods for satellite or reanalysis precipitation products [10–12]. These methods usually involve building a relationship between coarse-resolution precipitation data and high-resolution variables to improve the spatial resolution of satellite precipitation data. For example, Immerzeel et al. [13] established an exponential regression (ER) model by integrating 1-km normalized difference vegetation index (NDVI) data with Tropical Rainfall Measuring Mission (TRMM) 3B43 precipitation data and obtained high-resolution annual precipitation data over the Iberian Peninsula. Based on the method of Immerzeel, Jia et al. [14] established a functional relationship between 3B43 precipitation data and other variables (i.e., altitude and NDVI) using multiple linear regression (MLR) and obtained downscaled annual data at a 1 km resolution for the Qaidam Basin in China. Duan et al. [15] developed a further modified downscaling algorithm by introducing calibration methods based on geographic difference analysis (GDA) and geographic ratio analysis (GRA) and obtained 1-km monthly precipitation data over the Tana Lake Basin in Africa and the coast of the Caspian Sea in Asia. Zhang et al. [16] applied the abovementioned methods in Xinjiang, China, and obtained 1-km annual precipitation data. Considering the spatial variations exhibited by the relationship between precipitation and environmental variables, geographically weighted regression (GWR) has been introduced into precipitation analyses to achieve improved downscaling performance [17–20]. However, although environmental variables play a vital role in the monthly or yearly downscaling of precipitation, they have limited applicability in daily and hourly downscaling, which is more strongly reliant on cloud properties. For example, Sharifi et al. [21] obtained 1-km daily precipitation data in northeast Austria using MLR, artificial neural networks (ANNs) and spline interpolation based on 1-km cloud optical thickness (COT), cloud effective radius (CER), and cloud water path (CWP) data. Ma et al. [22] obtained 1-km hourly precipitation data in the southeast coast region of China based on COT, CER, and cloud top height (CTH) data from Himawari 8.

Traditional downscaling methods can improve the spatial resolution of satellite precipitation data. Many studies have shown that the accuracy of the satellite precipitation products is the most important factor affecting the quality of the downscaled estimates [13,14,17] if the environmental variables can satisfactorily reproduce the pattern of the satellite precipitation data. However, all previous downscaling methods have been applied to original satellite precipitation data, which contain large uncertainties that limit the accuracy of the downscaled precipitation estimates. Correcting the original satellite products before applying them in downscaling analyses can potentially contribute to the improvement of the downscaled precipitation estimates. Nevertheless, to our knowledge, no such study has previously been performed.

To address this gap, this study proposes a two-step merging and downscaling framework called OI-GWR to improve the accuracy of downscaled precipitation estimates. The merging procedure is based on optimum interpolation (OI), and the downscaling procedure is based on GWR. The remaining sections of this paper are organized as follows: Section 2 introduces an overview of the study area; Section 3 provides detailed information about the data and methods; Section 4 reports the results of OI and GWR; and finally, a discussion and conclusions are presented in Sections 5 and 6, respectively. The method proposed in this study will contribute to the production of high-quality and high-resolution regional gridded precipitation datasets. In particular, the method can serve as a useful reference for the development of grid data at the daily or hourly scale.

2. Study Area

The Tianshan Mountains are the mountain system that is the farthest in the world from any ocean. They are composed of a series of mountain ranges, intermountain basins, valleys, and piedmont plains, with an area of approximately 570,000 km^2 and an average altitude of 4000 m [4]. The precipitation distribution shows large temporal and spatial variations, with most precipitation occurring in summer and little occurring in winter. The Tianshan Mountains are also known as the water tower of central Asia. Approximately 65% of the rivers in Xinjiang originate from this region. On northern slopes, the annual average precipitation ranges between 500 and 700 mm, whereas the annual precipitation over western windward slopes can reach 1000 mm [23–25].

3. Data and Methodology

3.1. Data

3.1.1. Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) Satellite Precipitation Products

The Global Precipitation Measurement (GPM) mission is a satellite precipitation measurement project initiated by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). The goal of the GPM mission is to provide new-generation global satellite precipitation products with high precision and resolution. The GPM mission has an extended TRMM sensing load and an enhanced capacity for precipitation detection. The dual-frequency radar carried by the GPM Core Observatory (GPMCO) operates in the Ku and Ka bands. In particular, in the Ka band, it can operate in the high-sensitivity interleaved sampling mode. Meanwhile, the microwave radiometer of the GPMCO operates in four bands with higher frequencies than that of the TRMM Microwave Imager, which increases the observation capacity for light and solid precipitation. A comparison between the Integrated Multisatellite Retrievals for GPM (IMERG) products and other commonly used satellite precipitation datasets in Xinjiang has shown that IMERG exhibits the best performance [26–28]. Therefore, the IMERG monthly precipitation data were utilized as the initial data in this study.

IMERG can provide quasi-global precipitation data with a temporal resolution of 30 min and a spatial resolution of 0.1° × 0.1°. According to the calibration methods and data sources used for these precipitation data, the IMERG products can be divided into three types, namely, "Early-run", "Late-run", and "Final-run" products. Among them, the "Early-run" and "Late-run" products are quasi-real-time data in the sense that they are released 4 and 12 h after the observations, respectively, while the "Final-run" products are post-real-time data with a time lag of 3.5 months. These products are available at the Precipitation Measurement Missions (PMM) website (https://pmm.nasa.gov/data-access/downloads/gpm). Specifically, the V06B IMERG "Final-run" product was selected for use in this study because this product is subjected to gauge adjustment using monthly observation data from the Global Precipitation Climatology Centre (GPCC) and offers higher precision than the "Early-run" or "Late-run" products. There are only 9 Global Telecommunication System (GTS) stations used by the GPCC in the Tianshan Mountains; these stations account for a very small proportion (<1%) of the total AWSs in the region (1074) and were excluded in this study to ensure the independence of the precipitation evaluation. Because the AWSs in the study area cannot measure snowfall during the cold season, the study period was restricted to the warm seasons (May to September) from 2014 to 2018.

3.1.2. Observed Precipitation

Daily precipitation data were collected from 1065 AWSs in the Tianshan region and accumulated to the monthly scale. The distribution of the stations is shown in Figure 1. The time range of these precipitation observations was consistent with that of the IMERG precipitation data. To ensure the independence of model training and validation, the 9 GTS stations in the study area were excluded (Figure 1). The observed data were collected by the Information Center of the Xinjiang Meteorological

Bureau, and the quality control procedures included a climatic extreme value test, a single-station extreme value test, and a data consistency test. The data from September 2018 were excluded due to missing records for a large number of regional stations. In addition, since retrospective IMERG data are available from June 2000 to the present, a longer period of data was used to validate the effectiveness of the OI-GWR method in Section 5.2.
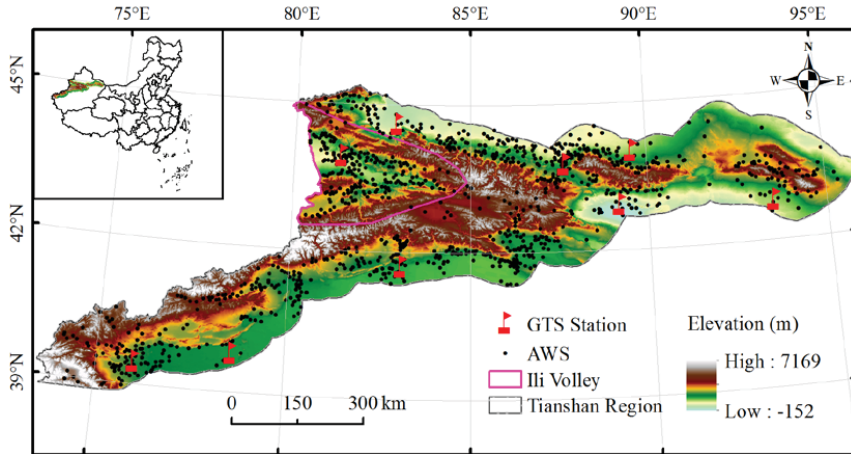
Figure 1. Digital elevation model (DEM) and distribution of the observation stations on a map of the Tianshan Mountains. The black dots and red flags represent automatic weather stations (AWSs) and Global Telecommunication System (GTS) stations, respectively.

3.1.3. Environmental Variables

The introduction of environmental variables into a downscaling analysis can improve the quality of satellite-retrieved precipitation data [13–20]. Therefore, environmental variables that are closely related to satellite precipitation data, such as the slope (SLP), aspect (ASP), curvature (CVT), hillshade (HSHD) [29,30], topographic wetness index (TWI) [31], and NDVI, were selected in this study. The NDVI data were obtained from the 1-km MOD13A3 monthly average vegetation index provided by the MODerate resolution Imaging Spectroradiometer (MODIS), and digital elevation model (DEM) data were taken from the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM), with a native resolution of 30 m. To maintain consistency with the NDVI data, the resolution was resampled to 1 km using the pixel averaging method, that is, the data from all high-resolution pixels within a given coarse-resolution pixel were averaged to obtain the corresponding coarse-resolution estimate. The topographic variables were obtained from the DEM data using Geographic Information System (GIS) software.

For the selected environmental variables, such as the NDVI, DEM, ASP, SLP, HSHD, TRI, and CVT, experiments were performed on each of the variables individually and on various combinations of variables. Moreover, the variance inflation factor (VIF) method was utilized to prevent multicollinearity. The VIF is a measure of the severity of multicollinearity in an MLR model, where multicollinearity refers to linear correlations between independent variables. The VIF is calculated when filtering variables. When the VIF value is closer to one, the multicollinearity is weaker, and vice versa. Specifically, when performing the experiments, variables were added to the existing variable group one by one. If adding a particular variable resulted in multicollinearity according to the VIF, that variable was deleted from the variable group. The variables were then introduced into the GWR model separately or in various combinations to obtain the downscaling outcomes. After a comparison of the results, the NDVI was found to have the best downscaling effect and thus was selected as the final explanatory variable for further study.

3.2. Methodology

3.2.1. Overall Flow Chart of the Study

In this study, monthly IMERG precipitation data were used for downscaling. As shown in Figure 2, first, the original IMERG precipitation data (OIMERG) and the observed precipitation data (OBS) were merged using OI to obtain 10-km OI-corrected IMERG precipitation data (CIMERG). The OI technique was originally developed by Eliassen [32] and Gandin [33]. In this study, OI was used to calibrate and adjust the satellite precipitation data based on OBS; hence, the corrected results possessed the combined advantages of satellite precipitation data and OBS (see Section 3.2.2 for details). Then, since it was found that the NDVI showed the best performance among all considered environmental variables (DEM, SLP, ASP, CVT, HSHD, TWI, and NDVI) in the downscaling of precipitation data, we included only the NDVI variable in this study, which was resampled from its original 1 km resolution to a 10 km resolution. With the 10-km NDVI as the independent variable, the spatial downscaling of CIMERG was performed using GWR to obtain a 1-km monthly precipitation dataset (DS_CIMERG) for the Tianshan Mountains. For comparison with previous methods, two additional sets of data were generated and used for further validation: DS_OIMERG, obtained via GWR downscaling with OIMERG (the original uncorrected IMERG precipitation data) as the dependent variable, and DS_Spline, obtained by using the spline interpolation method for OIMERG precipitation downscaling. The steps highlighted in the red dashed box in Figure 2 constitute the process of residual correction, which is necessary for the verification of the downscaling method proposed in this study.
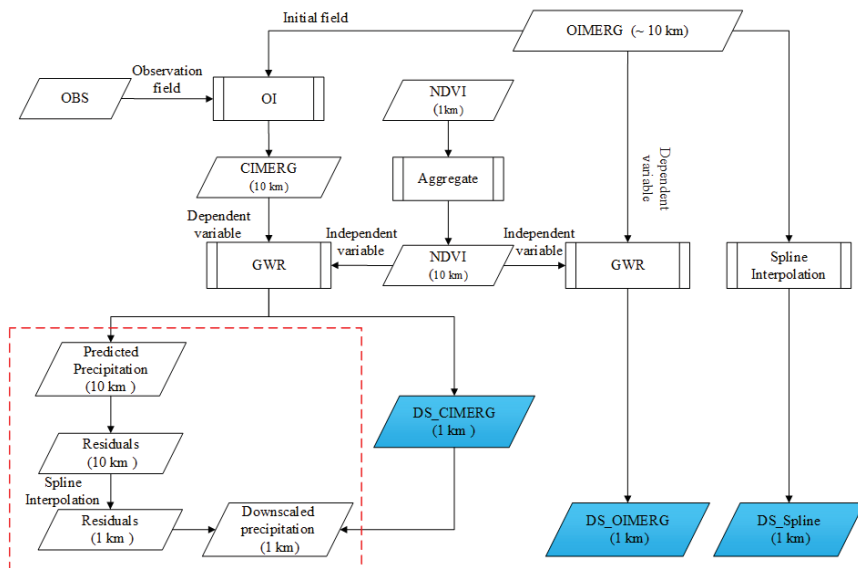
Figure 2. The flow chart of this study. The red dashed box indicates the residual correction process, and the blue boxes represent the three categories of downscaling estimation data. OIMERG, original Integrated Multi-satellitE Retrievals for GPM; OBS, observed precipitation data; OI, optimum interpolation; NDVI, normalized difference vegetation index; CIMERG, OI-corrected IMERG precipitation data; GWR, geographically weighted regression; DS_CIMERG, obtained via GWR downscaling with CIMERG; DS_OIMERG, obtained via GWR downscaling with OIMERG; DS_Spline, obtained by using the spline interpolation method for OIMERG precipitation downscaling.

3.2.2. Optimum Interpolation (OI)

OI analysis requires an initial estimation field, such as a set of gridded satellite precipitation data. By calculating the error weight function of the initial estimation field and the observation field point by point, the target grid points for analysis can be corrected. Thus, when developing OI-based data merging algorithms, the key is to quantify the error structure of the initial estimation field and the observation field. Unlike other data-merging methods, OI considers not only the autocorrelation of various errors but also the correlation between different observations. Moreover, OI involves solving for the optimal value within a certain range of each analysis point and thus is particularly suitable for the analysis of single variables with large spatiotemporal variability, such as precipitation [34]. In the OI analysis conducted in this study, the OIMERG precipitation data were used as the initial estimation field, and the station precipitation observations were used as the observation field. The final analysis result for the precipitation value (A_k) at each grid point is equivalent to the first guess (F_k) at this grid point plus the deviation between the observed value and the initial estimated value at the grid point. This deviation is obtained through weighted estimation based on the deviations between the known observed values (O_i) and initial estimated values (F_i) from n grid points within a certain range, and it represents the maximum distance correlated with the target grid point. The formula is as follows:

$$A_k = F_k + \sum_{i=1}^{n} W_i (O_i - F_i) \tag{1}$$

where k is the grid point to be analyzed, i is an index representing the "valid grid boxes" (boxes in the satellite precipitation grid containing at least one gauge station), and W_i is a weight coefficient assigned to the deviation between the observed value and the initial estimated value in the ith grid box during estimation. Note that in an area with a sparse station network, the analysis radius should be continuously adjusted to ensure that a sufficient number of valid grid boxes can be searched, from which several valid grid boxes nearest to the target grid point are then selected for inclusion during OI. In this study, the analysis radius was set to 100 km, and the 9 nearest valid grid boxes to the target grid point were selected for OI [35].

In Equation (1), the weight coefficients (W_i) are determined by the variance in the minimum error on the precipitation value (A_k) at the target point:

$$E^2 = \overline{(A_K - T_K)^2} \tag{2}$$

where T_K represents the observed value at point k.

Based on the assumptions that the observation field and the initial estimation field are unbiased and that the observation error is not related to the error of the initial estimation field, the weight coefficients W_i in Equation (1) can be obtained by solving the following linear equation [34]:

$$\sum_{j=1}^{n} \left(\mu_{ij}^f + \mu_{ij}^o \lambda_i \lambda_j \right) W_j = \mu_{ki}^f \tag{3}$$

where μ_{ij}^f represents the co-correlation of the initial estimation field error, μ_{ij}^0 represents the co-correlation of the observation error, and λ_i is the ratio between the standard deviation of the observation error (σ_i^o) and that of the initial estimation error (μ_{ij}^o) at point i. In OI, the calculation of the W_i requires that μ_{ij}^f, μ_{ij}^o, σ_i^o, and σ_i^f are known values, which, in turn, requires the pre-estimation of the observation error and the satellite-retrieved precipitation error as well as the correlation between these errors. Here, the term "pre-estimation" means that these parameters need to be estimated in advance. In this study, this pre-estimation was performed based on a statistical analysis of the sample data within the study period [35].

The weight coefficients (W_i) were determined based on Equation (3), and then, the final precipitation values (A_k) were obtained based on Equation (1).

3.2.3. GWR Downscaling

Since the GWR model was first proposed by Brunsdon et al. [36] in 1996, it has been extensively applied in research on spatial heterogeneity [17–20,37,38]. The basic idea of GWR is that the relationship between variables varies with changes in spatial location; thus, a regression model can be established by estimating the parameters of the correlated variables and explanatory variables at each given location in the study area. Figure 3 shows the spatial distributions of the intercept, NDVI regression coefficient, and local R^2 obtained via GWR, and these values are in accordance with the definition. These parameters exhibit significant spatial variations. The intercept ranges from -78.7 to 178.4, the NDVI coefficient ranges from -123.3 to 203.4, and the local R^2 ranges from 0.01 to 0.92.
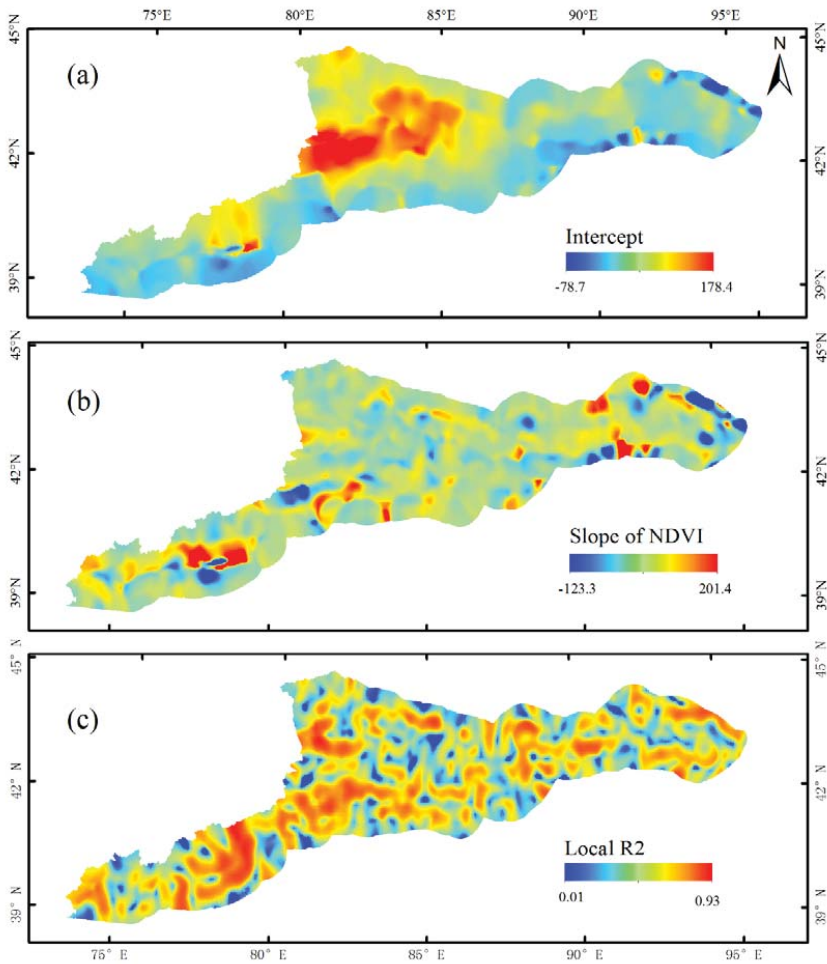
Figure 3. The spatial distributions of the (**a**) intercept, (**b**) slope of the NDVI and, (**c**) local R^2 in July 2016.

In this study, a GWR regression model was established based on the NDVI and the IMERG precipitation data as follows:

$$Y_j = \beta_0(u_j, v_j) + \sum_{i=1}^{p} \beta_i(u_j, v_j)X_{ij} + \varepsilon_j \tag{4}$$

where Y_j is the IMERG precipitation at point j; X_{ij} is the NDVI at point i in the vicinity of point j; $\beta_0(u_j, v_j)$ and $\beta_i(u_j, v_j)$ represent the intercept and slope, respectively, at point j; (u_j, v_j) represents the two-dimensional coordinates of point j; and ε_j is the residual error. Unlike traditional global regression models, Equation (4) is based on the assumption that the shorter the distance between the observation point and point j is, the greater the influence on point j will be, with the coefficient acting as a damping function that depends on the distance from point j. This damping function can be obtained in accordance with Equation (5):

$$\hat{\beta}(u_j, v_j) = \left(X^T\left(W(u_j, v_j)\right)X\right)^{-1}X^T W(u_j, v_j)Y \tag{5}$$

where $\hat{\beta}(u_j, v_j)$ represents the coefficient of point j; X and Y are the independent and dependent variables, respectively; and $W(u_j, v_j)$ is a weight matrix. This matrix ensures that the shorter the distance between points i and j is, the greater the weight, and the elements of the matrix can be obtained as follows:

$$w_{ij} = \left[1 - \left(d_{ij}/b\right)^2\right]^2 \text{ when } d_{ij} \leq b$$
$$w_{ij} = 0 \text{ when } d_{ij} > b \tag{6}$$

where d_{ij} is the distance of point j from the nearby observation point i, and b is a fixed threshold defined in terms of a distance metric.

In detail, the following procedures were applied for GWR-based downscaling (Figure 2).

(1) To effectively establish the precipitation-NDVI model, anomalous NDVI areas corresponding to snow and water bodies were removed from the high-spatial-resolution NDVI data [17,39].

(2) After the removal of outliers, the 1-km NDVI data were aggregated to a resolution of 10 km by means of pixel averaging. Then, a GWR model of the 10-km IMERG data and the 10-km NDVI data was established with the NDVI as the independent variable and the IMERG precipitation data as the dependent variable. By introducing 1-km and 10-km grid point coordinates into the GWR model, the constants and corresponding coefficients for the 1-km and 10-km NDVI were obtained, as shown in Equation (4).

(3) The 10-km NDVI data were entered into the regression model to obtain NDVI-based precipitation predictions with a 10 km resolution (Predicted Precipitation 10 km in Figure 2).

(4) The residual errors between the values predicted by the 10-km resolution model and the original IMERG precipitation values were calculated (Residuals 10 km in Figure 2). The 10-km residuals were then transformed into 1-km residuals through spline interpolation (Residuals 1 km in Figure 2).

(5) The 1-km NDVI data after anomalous data removal were used to force the regression model, thus obtaining 1-km model-predicted precipitation values. Spline interpolation was then applied to fill in the values missing after outlier removal to obtain downscaled data (DS_CIMERG in Figure 2).

(6) The 1-km model-predicted precipitation values and 1-km residual data were summed to obtain post-residual-corrected 1-km downscaled precipitation data.

3.2.4. Validation

Station-measured data are the most direct observations of precipitation. In this study, 10-fold cross-validation was used to validate the precision of the OI outcomes and downscaling products. The observation stations were randomly divided into 10 groups. Nine groups (90%) were selected for the OI-based merging of the IMERG precipitation data with the observation data and GWR-based

downscaling. The remaining 10% constituted an independent dataset to validate the accuracy of the OI results and downscaling products. This was repeated 10 times to guarantee an independent validation for each station and to ensure the representativeness of the training samples and validation samples. Three statistical indicators, namely, the mean absolute error (MAE), root-mean-square error (RMSE), and correlation coefficient (CC), were utilized to validate the estimated values of the downscaled product against the observed values. The formulas for these indicators are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{8}$$

$$CC = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{9}$$

where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$, $\overline{y} = \frac{1}{n}\sum_{i=1}^{n}y_i$, n is the sample size, and x_i and y_i are the estimated values and station observations of precipitation, respectively.

In addition, a statistical analysis showed that July 2016 was the month with the most precipitation in the Tianshan Mountain area, while September 2017 was the month with the least precipitation. Therefore, in addition to the 10-fold cross-validation, the data from these two months were used to further validate the results of the proposed downscaling method in terms of their spatial distribution.

4. Results

4.1. Consistency Between the NDVI and Satellite Precipitation Data

To determine the delay in the response of the NDVI to precipitation, Table 1 presents the correlation coefficients (CCs) between the NDVI and the observed precipitation at all stations in the study area during the warm seasons from 2014 to 2018. The results showed that the highest correlations were observed with no time lag. Figure 4 shows the distributions of the monthly average NDVI and IMERG precipitation in the study area during the warm season of 2016. According to Figure 4b, the Ili Valley was the area with the most abundant precipitation, while the eastern area received the least precipitation. This is because the Ili Valley is surrounded by mountains to the north, east, and south (Figure 1), which is beneficial for the collection of water vapor from the Atlantic Ocean to the west, resulting in abundant precipitation in the piedmont zone. According to Figure 4a, there was good consistency between the NDVI and IMERG precipitation data. The high resolution of the NDVI revealed detailed spatial variations and contributed to the improvement of the downscaled precipitation estimates.

Table 1. The correlation coefficients (CCs) with different delays between the NDVI and the observed precipitation at all stations during the warm seasons from 2014 to 2018.

Lag (months)	0	1	2	3
CC	0.56	0.50	0.39	0.27
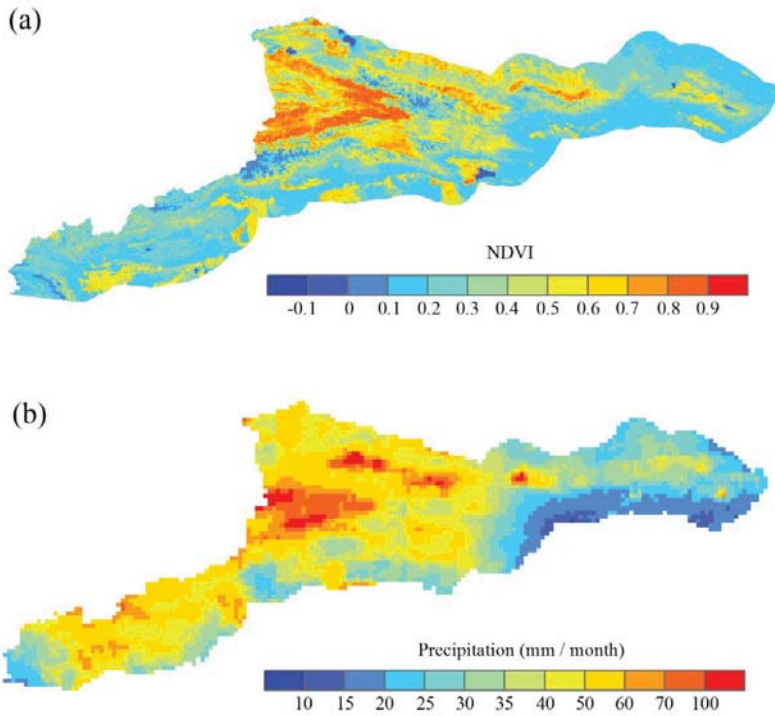
Figure 4. (**a**) The monthly average NDVI and (**b**) the monthly average IMERG during the warm season of 2016.

4.2. Improvement of the OIMERG Precipitation Data Achieved via OI

Figure 5 shows comparisons of the CC, MAE, and RMSE values calculated before and after the OI-based merging of the OIMERG data and the observation data. After OI, the CC values increased in almost all months; this effect was the most obvious for September 2017, with an increase of 0.27. The CC was slightly worse after OI for only two months: by 0.01 for May 2014 and by 0.02 for August 2014. In addition, the RMSE generally decreased by 0.37–5.16 mm, with slight increases observed for only two months (by 0.06 mm for September 2016 and by 0.23 mm for August 2014). The MAE decreased in all months except August 2014, for which it increased by 0.17 mm. Overall, compared with the OIMERG precipitation data, the CIMERG data obtained after OI showed improvements in the mean CC, RMSE, and MAE values from 0.516, 28.83 mm, and 19.73 mm to 0.616, 26.56 mm, and 17.59 mm, respectively. Therefore, the OI processing of the OIMERG and observed precipitation data greatly improved the precision of the original satellite precipitation product.
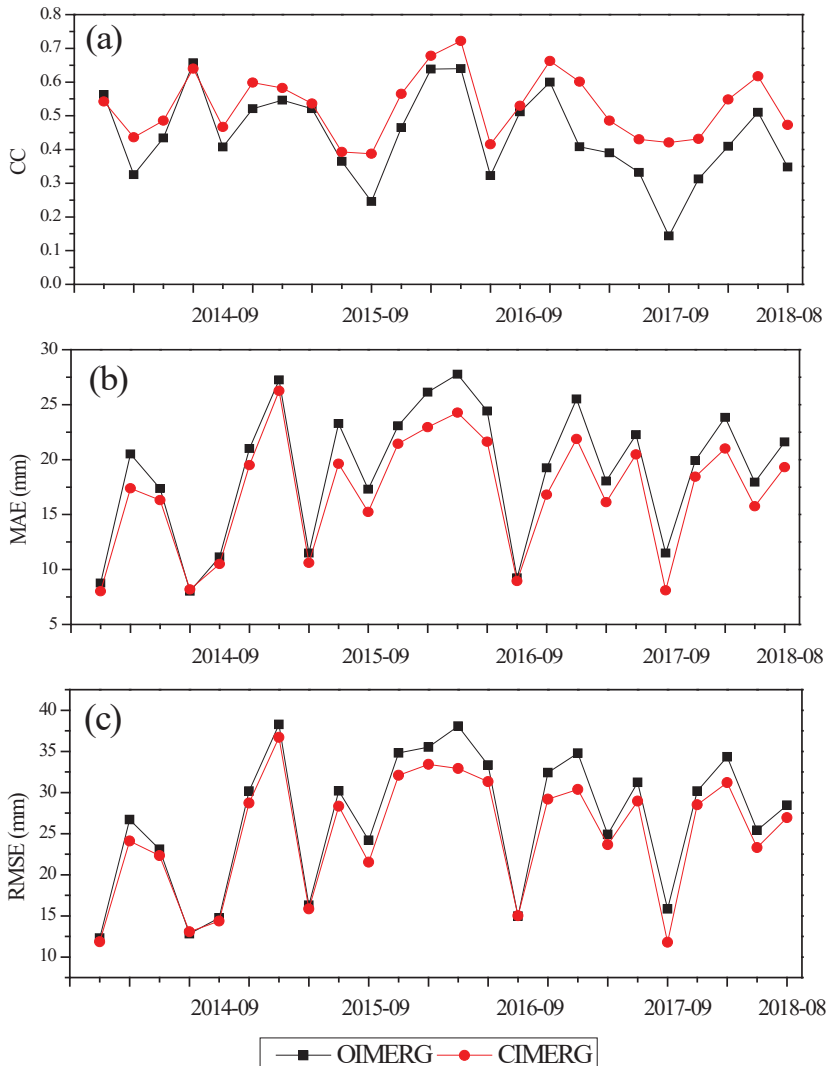
Figure 5. Comparisons of the (**a**) CC, (**b**) mean absolute error (MAE) and (**c**) root-mean-square error (RMSE) before and after OI processing for the warm seasons of 2014–2018.

4.3. Error Statistics of DS_Spline, DS_OIMERG, and DS_CIMERG

Based on the improvement in the quality of the satellite precipitation product after OI, the downscaling effect of the GWR method was further validated. Figure 6 shows comparisons among the downscaled data obtained with three different downscaling methods for May to September of 2014–2018. DS_CIMERG had the highest CC values in 22 of the 24 months, with lower values than those of DS_Spline only in May and August 2014. On average, the mean CC of DS_CIMERG was 10% higher than that of DS_Spline. Between DS_Spline and DS_OIMERG, the CC values of DS_Spline were higher in 8 months and lower in 16 months, suggesting that the quality of satellite precipitation data cannot be effectively improved by downscaling with only the spatial interpolation technique. The RMSE and MAE values of DS_CIMERG were the lowest

in almost all months. Furthermore, the mean CC, RMSE, and MAE values respectively improved from 0.616, 26.56 mm, and 17.59 mm for CIMERG to 0.635, 25.93 mm, and 17.15 mm for DS_CIMERG, respectively.
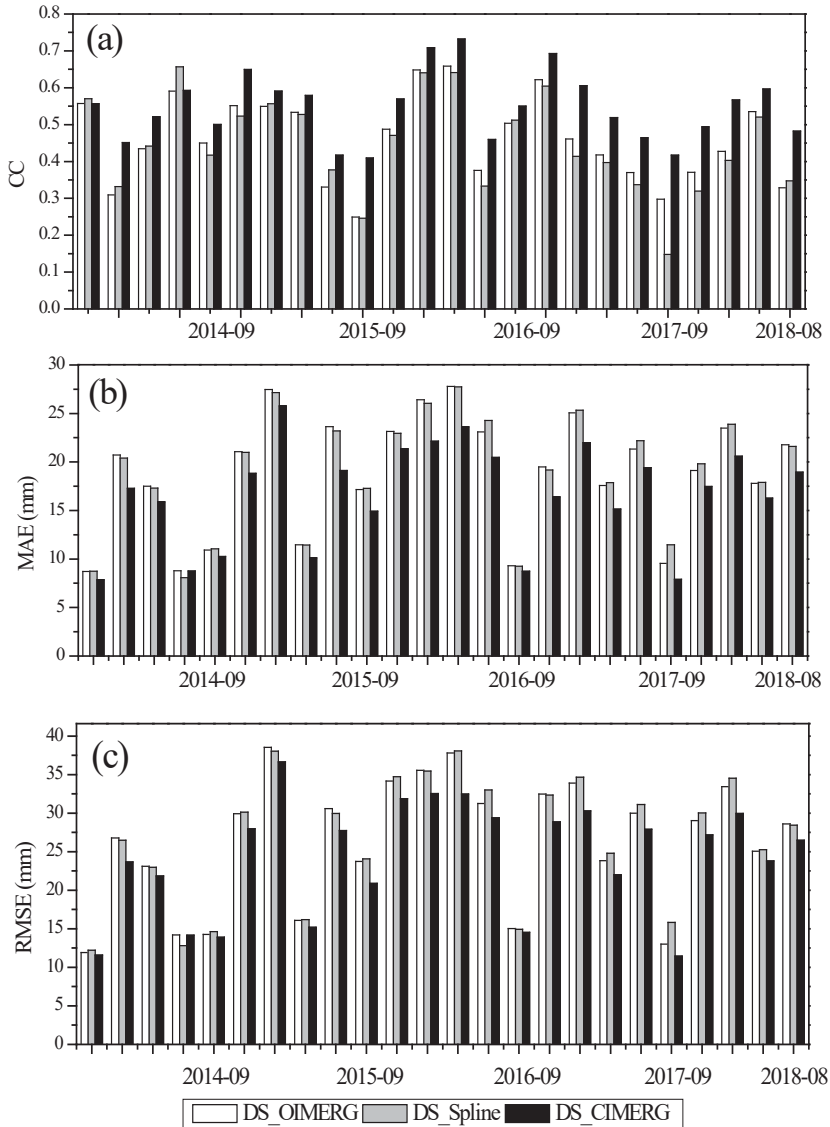
Figure 6. (**a**) The correlation coefficient (CC), (**b**) mean absolute error (MAE), and (**c**) root-mean-square error (RMSE) values of downscaled precipitation estimates obtained using GWR based on OIMERG (DS_OIMERG), GWR based on CIMERG (DS_CIMERG), and spline interpolation (DS_Spline) for the Tianshan Mountains between May and September 2014–2018. Independent rain gauge observations were used as benchmarks.

4.4. Overall Assessment of the Two-Step Merging and Downscaling Method

Figure 7 presents scatter plots of OIMERG, CIMERG, and the three types of downscaled precipitation data against all of the observation data from the study area. After the first and second processing steps (i.e., OI and GWR), the CC was improved by 10% and 2%, respectively. As shown in Figure 7e, DS_OIMERG showed slight improvements compared with OIMERG, while it was less accurate than CIMERG and DS_CIMERG, indicating that direct downscaling based on original satellite precipitation products may have a limited effect in improving the quality of downscaled precipitation estimates.
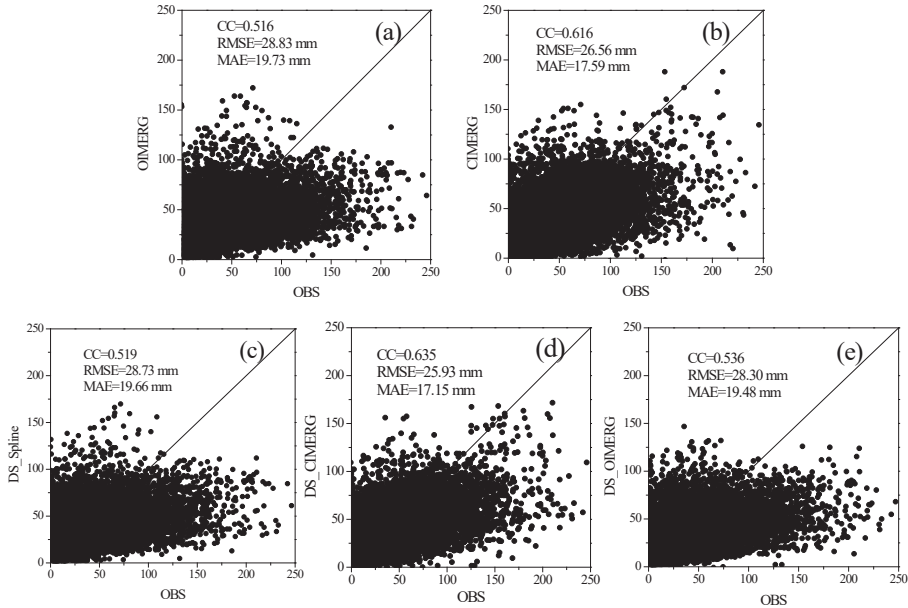

Figure 7. Scatter diagrams of the (**a**) OIMERG, (**b**) CIMERG, (**c**) DS_Spline, (**d**) DS_CIMERG, and (**e**) DS_OIMERG precipitation products against the observed precipitation during the warm seasons from 2014 to 2018. OIMERG, original IMERG precipitation data; CIMERG, OI-corrected IMERG precipitation data; DS_Spline, obtained by using the spline interpolation method for OIMERG precipitation downscaling; DS_CIMERG, obtained via GWR downscaling with CIMERG; DS_OIMERG, obtained via GWR downscaling with OIMERG; OBS, observed precipitation data.

Figure 8 shows boxplots of the evaluation metrics for the five precipitation products. A boxplot divides a dataset into four segments based on the maximum, minimum, median, and two quartiles of the data. The middle horizontal line represents the median, which divides the statistical data into two equal parts. As shown in Figure 8, the distributions of the CC, MAE, and RMSE were all uniform and consistent. Among the five datasets, the best performance was observed for DS_CIMERG, whose CC values were more concentrated in the upper region, while the MAE and RMSE values were more concentrated in the lower region. In particular, the minimum and maximum MAE and RMSE were the lowest for DS_CIMERG, followed by CIMERG. The metrics for OIMERG and DS_Spline were basically the same, while those for DS_OIMERG were in agreement with the overall assessment results and slightly better than those for OIMERG (Figure 7). In Figure 8, the small rectangular boxes represent the distributions of the metric values around their averages. DS_CIMERG performed the best among the five products, with CC, MAE, and RMSE values of 0.70, 17.29 mm, and 22.45 mm, respectively. The overall evaluation results showed that the dataset generated using the OI-GWR method proposed

in this study was notably superior to the datasets of the other two downscaling products as well as the initial IMERG precipitation data.
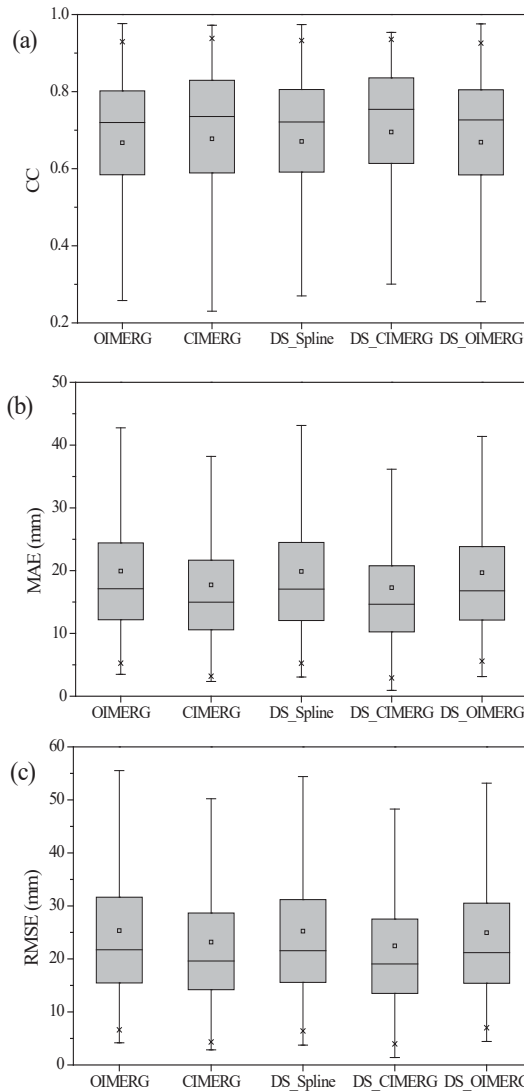

Figure 8. Boxplots of the evaluation metrics (**a**, CC; **b**, MAE and **c**, RMSE) for the five precipitation datasets for the warm seasons of 2014–2018. OIMERG, original IMERG precipitation data; CIMERG, OI-corrected IMERG precipitation data; DS_Spline, obtained by using the spline interpolation method for OIMERG precipitation downscaling; DS_CIMERG, obtained via GWR downscaling with CIMERG; DS_OIMERG, obtained via GWR downscaling with OIMERG.
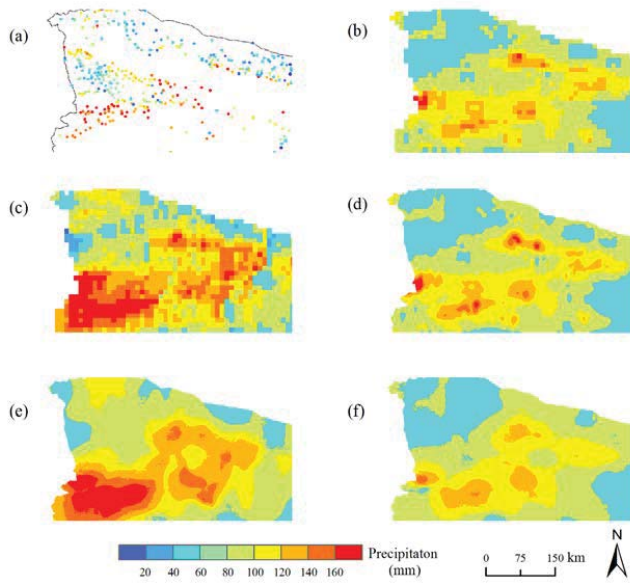
4.5. Spatial Distributions of Monthly Precipitation

Using July 2016 (the month with the maximum precipitation) and September 2017 (the month with the minimum precipitation) as examples (Figures 9–12), the spatial distributions of the OBS, OIMERG, CIMERG, DS_OIMERG, DS_Spline, and DS_CIMERG monthly precipitation data are shown in Figure 9;

Figure 11. Furthermore, Figure 10; Figure 12 show the detailed distributions of the precipitation datasets in the Ili Valley after image enlargement. Figure 9a shows the observed precipitation distribution in July 2016. The Ili Valley was the region with the most abundant precipitation. The other five gridded estimated precipitation datasets (OIMERG (Figure 9b), CIMERG (Figure 9c), DS_Spline (Figure 9d), DS_CIMERG (Figure 9e), and DS_OIMERG (Figure 9f)) all exhibited the same spatial distribution pattern as that of the OBS dataset. However, OIMERG underestimated the observed precipitation, particularly in the Ili Valley area. In contrast, CIMERG, obtained by combining OIMERG and ground observations using OI, showed a more reasonable precipitation distribution. Figure 10 shows a clearer view of the patterns in this area. Since DS_Spline and DS_OIMERG were generated by downscaling the original IMERG data, their performance was limited by the IMERG precipitation distribution, which underestimated the observed precipitation. By contrast, after OI-GWR processing, DS_CIMERG not only showed noticeably improved accuracy relative to the initial IMERG precipitation data but also exhibited a precipitation distribution that was more detailed and consistent with the observed precipitation, thus demonstrating the improved capabilities of the proposed downscaling method. Figure 11 presents the spatial distributions of the precipitation data for September 2017 (the month with the lowest precipitation). It can be seen that, during this time period, the precipitation was weak throughout the Tianshan Mountain area (Figure 11a). Consistent with the results shown in Figure 9, the DS_CIMERG dataset showed a distribution more similar to the observed precipitation distribution (Figure 11e), while the distributions of DS_Spline (Figure 11d) and DS_OIMERG (Figure 11f) were consistent with that of OIMERG.
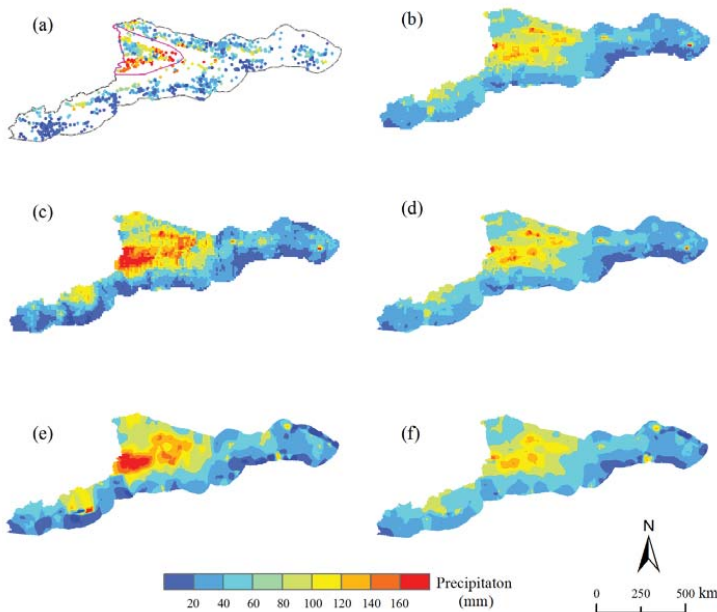
Figure 9. The spatial distributions of the five estimated precipitation datasets and the observed precipitation in July 2016. (**a**) OBS; (**b**) OIMERG; (**c**) CIMERG; (**d**) DS_Spline; (**e**) DS_CIMERG; (**f**) DS_OIMERG.

Figure 10. Enlarged views of the Ili Valley corresponding to Figure 9. (**a**) OBS; (**b**) OIMERG; (**c**) CIMERG; (**d**) DS_Spline; (**e**) DS_CIMERG; (**f**) DS_OIMERG.
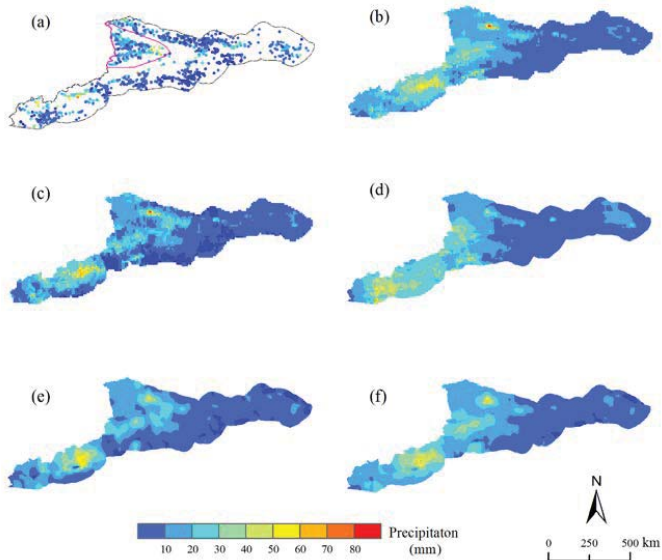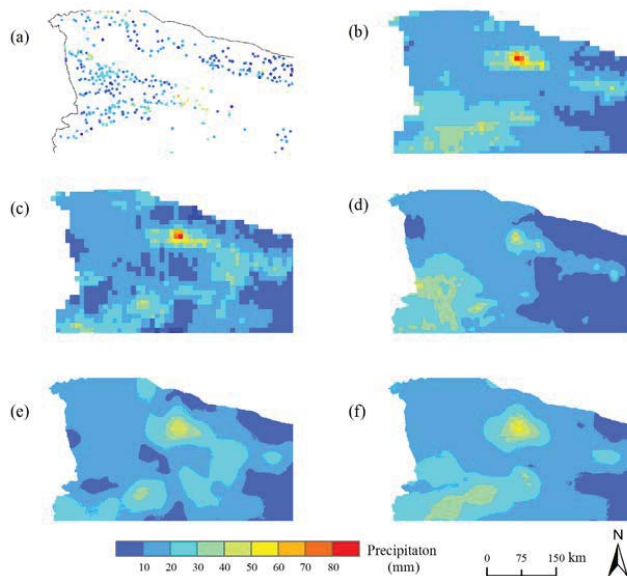

Figure 11. The spatial distributions of the five estimated precipitation datasets and the observed precipitation in September 2017. (**a**) OBS; (**b**) OIMERG; (**c**) CIMERG; (**d**) DS_Spline; (**e**) DS_CIMERG; (**f**) DS_OIMERG.

Figure 12. Enlarged views of the Ili Valley corresponding to Figure 11. (**a**) OBS; (**b**) OIMERG; (**c**) CIMERG; (**d**) DS_Spline; (**e**) DS_CIMERG; (**f**) DS_OIMERG.
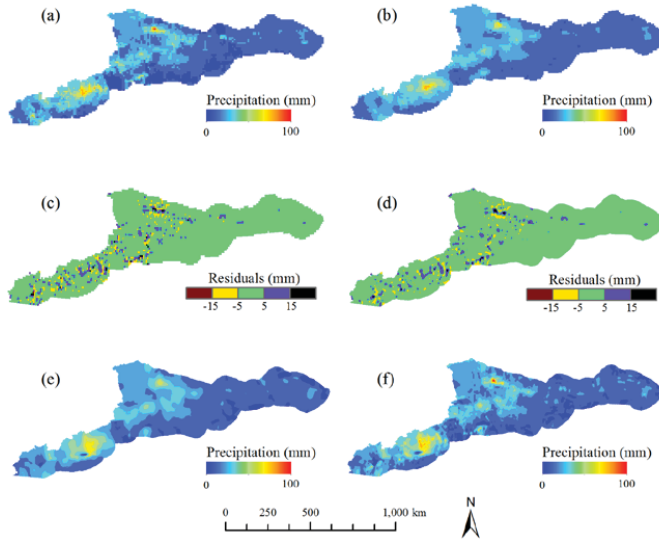
4.6. Residual Correction

A series of improvements in downscaling methods for satellite precipitation products have been achieved, from the original ER model proposed by Immerzeel et al. [13] and the MLR model proposed by Jia et al. [14] to the GWR approach proposed by Xu et al. [17] and Chen et al. [20]. Residual correction is a key step in both ER and MLR [17]. In this study, residual correction for GWR was analyzed in detail.

Using DS_CIMERG as an example, the residual correction processes for GWR downscaling in July 2016 and September 2017 were analyzed, and the results are shown in Figure 13; Figure 14, respectively. Figure 13a shows the CIMERG data before downscaling. Figure 13b shows the estimated precipitation based on the low-resolution (10 km) regression coefficients (i.e., Predicted Precipitation (10 km) in Figure 2). Figure 13c shows the 10-km residuals obtained by subtracting CIMERG from the 10-km precipitation predictions. Figure 13d shows the 1-km residuals obtained after applying the spline interpolation technique to the 10-km residuals. Figure 13e shows the DS_CIMERG precipitation estimates, which are based on the high-resolution (1 km) regression coefficients. Finally, Figure 13f presents the residual-corrected downscaling results obtained by summing the 1-km residuals and the 1-km precipitation estimates. As shown in this figure, the precipitation estimates (Figure 13b) obtained with the low-resolution parameters already exhibited high consistency with CIMERG (Figure 13a), and their residual values were generally small, ranging between -10 and 10 mm (Figure 13c), indicating that the low-resolution estimated precipitation data obtained via the GWR method were close to the initial precipitation data. In contrast, the high-resolution precipitation estimates obtained based on the 1-km regression coefficients (Figure 13e) not only were highly consistent with the initial precipitation distribution but also reflected the detailed structure of the precipitation distribution. The results shown in Figure 14 are similar to those in Figure 13, except that they correspond to a different month. Table 2 summarizes the detailed error statistics. The CC, RMSE, and MAE values after residual correction were worse than those before residual correction. Therefore, the addition of residual correction to the GWR model transferred the errors of the original IMERG precipitation data to the final downscaling outcomes, thereby decreasing the reliability of the outcomes, as the residuals were obtained by subtracting the 10-km precipitation predictions from the original IMERG precipitation data.
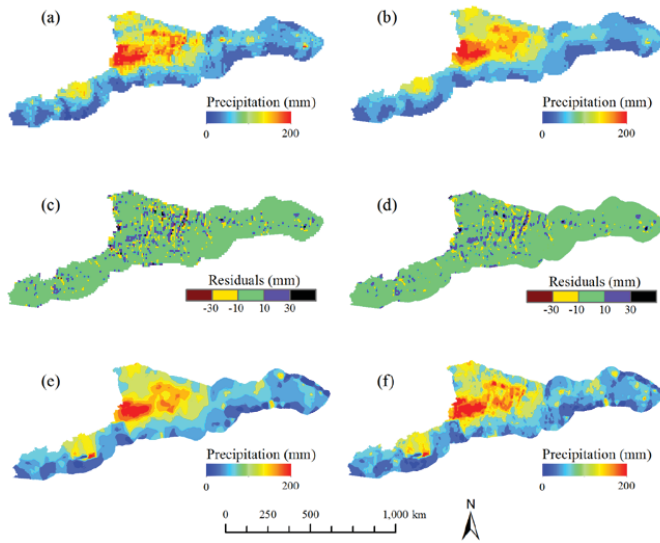
Figure 13. Comparison of the DS_CIMERG product in July 2016 before and after residual correction. (**a**) CIMERG; (**b**) predicted precipitation (10 km); (**c**) residuals (10 km); (**d**) residuals (1 km); (**e**) DS_CIMERG; (**f**) downscaled precipitation after residual correction.

Figure 14. Comparison of the DS_CIMERG product in September 2017 before and after residual correction. (**a**) CIMERG; (**b**) predicted precipitation (10 km); (**c**) residuals (10 km); (**d**) residuals (1 km); (**e**) DS_CIMERG; (**f**) downscaled precipitation after residual correction.

To further evaluate the results through additional validation tests, the data before and after residual correction from 2010 to 2018 were also processed and evaluated. The results were consistent with those presented in Table 2. The CC, RMSE, and MAE values respectively increased from 0.576, 26.92 mm, and 18.12 mm before residual correction to 0.59, 26.55 mm, and 17.63 mm after residual

correction. Therefore, the residual correction step is not necessary in GWR downscaling; this conclusion is consistent with the findings of Xu et al. [17].

Table 2. Comparison of evaluation metrics before and after residual correction.

| Time	Residual Correction	CC	RMSE (mm)	MAE (mm)
07/2016	Before	0.718	33.03	23.82
	After	0.695	34.11	24.59
09/2017	Before	0.435	11.54	7.80
	After	0.425	11.90	7.91

5. Discussion

5.1. Limitations of GWR Downscaling

Previous studies have shown that, in the case that an established model can adequately estimate precipitation, the quality of the initial satellite precipitation data becomes a critical factor in determining the quality of the downscaling results [14,17]. In this study, the proposed two-step OI-GWR method combining data merging and downscaling successfully solved this problem. Nevertheless, some limitations of the GWR method were also found during this research. First, during the process of variable selection for GWR, when the variables with the best global correlation with precipitation were applied for GWR downscaling, the obtained results were not optimal. In addition, the inclusion of more explanatory variables (e.g., DEM, SLP, ASP, CVT, HSHD, and/or TWI) did not yield better results than those achieved using the single NDVI variable. Second, unlike the stepwise regression method, GWR cannot automatically identify and eliminate variables that are not significantly related to the dependent variables, which can be considered a shortcoming of the GWR method itself. Even after stepwise regression was performed for variable screening and the selected variables were input into the GWR model, the desired results were not obtained. Presumably, this can be explained by the fact that stepwise regression is a global regression method, which therefore selects variables from the perspective of global correlation, whereas the GWR method relies on point-by-point regression and therefore focuses on local modeling. For this reason, variables selected on the basis of stepwise regression or correlation analysis may not be suitable for GWR. Accordingly, further investigation will be necessary to improve the process of variable selection for GWR.

5.2. Impact of the Gauge Density on OI-GWR

To further validate the performance of OI-GWR and its applicability in sparsely gauged regions, we randomly selected 30% of the rain gauges for model training and used the remaining 70% of the rain gauges for model validation. For this analysis, the study period was extended to 2010-2018. Figure 15 shows that even when using only 30% of the stations for training, a positive correction effect could be obtained. However, since fewer rain gauges were used for training, the improvement was less significant than that achieved using more rain gauges. An overall evaluation showed that after the two steps of correction (i.e., OI and GWR), the CC increased from the initial value of 0.533 to values of 0.567 and 0.59, respectively; the RMSE decreased from the initial value of 27.08 mm to values of 26.84 and 26.55 mm, respectively; and the MAE decreased from the initial value of 18.3 mm to values of 17.76 and 17.63 mm, respectively. Therefore, the OI-GWR method proposed in this study can improve the deficiencies of the previous downscaling method. Since OI-GWR works well even when the size of the training dataset is reduced to 30%, the proposed method has great potential for application in sparsely gauged regions.
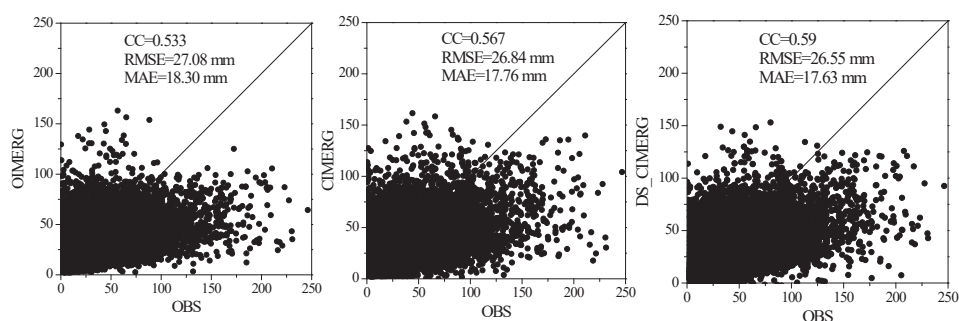
Figure 15. Scatter diagrams of OIMERG, CIMERG, and DS_CIMERG against the observed precipitation during the warm seasons from 2010 to 2018.

6. Conclusions

In this study, a two-step merging and downscaling method (OI-GWR) was proposed and used to downscale original IMERG precipitation data for the Tianshan Mountains from a resolution of ~10 to 1 km. First, the original IMERG precipitation data were merged with observed precipitation data using the OI method (https://github.com/tgq14/GWR-OI). Then, downscaling was performed using the GWR method, with the corrected CIMERG precipitation data as the initial values and the NDVI as an auxiliary variable. The performance of OI-GWR was assessed based on rain gauge observations, and the results were compared with those of other downscaling methods. The main conclusions include the following:

(1) The OI-based merging of the OIMERG data and the observed precipitation data greatly improved the precision of the original satellite precipitation product. After OI, the CC, RMSE, and MAE values for the OIMERG data were increased from 0.516, 28.83 mm, and 19.73 mm to 0.616, 26.56 mm, and 17.59 mm, respectively.

(2) An assessment of various downscaled datasets showed that the precision of DS_CIMERG (based on OI-GWR) was much higher than that of DS_OIMERG (based on GWR), indicating that the downscaling results for satellite precipitation data depend to a large extent on the quality of the initial precipitation product.

(3) Residual correction is a key step in global regression downscaling methods, such as ER and MLR. However, in the GWR method, residual correction tends to transfer the errors of the original satellite precipitation data to the final downscaling results. The statistical evaluation metrics of the GWR results obtained after residual correction were worse than those before residual correction, indicating that residual correction is unnecessary in GWR-based downscaling.

By improving the precision of the original IMERG satellite precipitation products and then downscaling the products thus obtained, the two-step OI-GWR method can serve as a more effective approach for the generation of regional precipitation datasets. Further efforts will be needed to extend the application of OI-GWR to the daily or hourly scale.

References

1. Sorooshian, S. Commentary-GEWEX (Global Energy and Water Cycle Experiment) at the 2004 Joint Scientific Committee Meeting. *GEWEX Newsl.* 14 Ferbuary 2004.
2. Ebert, E.E.; Janowiak, J.E.; Kidd, C. Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 47–64. [CrossRef]
3. Verdin, A.; Rajagopalan, B.; Kleiber, W.; Funk, C. A Bayesian kriging approach for blending satellite and ground precipitation observations. *Water Resour. Res.* **2015**, *51*, 908–921. [CrossRef]
4. Gao, L.; Wei, J.; Wang, L.; Bernhardt, M.; Schulz, K.; Chen, X. A high-resolution air temperature data set for the Chinese Tian Shan in 1979–2016. *Earth Syst. Sci. Data* **2018**, *10*, 2097–2114. [CrossRef]
5. Adhikary, S.K.; Yilmaz, A.G.; Muttil, N. Optimal design of rain gauge network in the Middle Yarra River catchment, Australia. *Hydrol. Process.* **2015**, *29*, 2582–2599. [CrossRef]
6. AghaKouchak, A.; Behrangi, A.; Sorooshian, S.; Hsu, K.; Amitai, E. Evaluation of satellite-retrieved extreme precipitation rates across the central United States. *J. Geophys. Res. Space Phys.* **2011**, *116*, 3–25. [CrossRef]
7. Chen, Y.; Ebert, E.E.; Walsh, K.J.; Davidson, N.E. Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN data. *J. Geophys. Res. Atmos.* **2013**, *118*, 2184–2196. [CrossRef]
8. Kidd, C.; Kniveton, D.R.; Todd, M.C.; Bellerby, T.J. Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms. *J. Hydrometeorol.* **2003**, *4*, 1088–1104. [CrossRef]
9. Xie, P.; Arkin, P.A. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *B. Am. Meteorol. Soc.* **1997**, *78*, 2539–2558. [CrossRef]
10. Jing, W.; Yang, Y.; Yue, X.; Zhao, X. A Spatial Downscaling Algorithm for Satellite-Based Precipitation over the Tibetan Plateau Based on NDVI, DEM, and Land Surface Temperature. *Remote Sens.* **2016**, *8*, 655. [CrossRef]
11. Wang, L.; Chen, R.; Han, C.; Yang, Y.; Liu, J.; Liu, Z.; Wang, X.; Liu, G.; Guo, S. An Improved Spatial–Temporal Downscaling Method for TRMM Precipitation Datasets in Alpine Regions: A Case Study in Northwestern China's Qilian Mountains. *Remote Sens.* **2019**, *11*, 870. [CrossRef]
12. Long, Y.; Zhang, Y.; Ma, Q. A Merging Framework for Rainfall Estimation at High Spatiotemporal Resolution for Distributed Hydrological Modeling in a Data-Scarce Area. *Remote Sens.* **2016**, *8*, 599. [CrossRef]
13. Immerzeel, W.; Rutten, M.; Droogers, P. Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sens. Environ.* **2009**, *113*, 362–370. [CrossRef]
14. Jia, S.; Zhu, W.; Lú, A.; Yan, T. A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sens. Environ.* **2011**, *115*, 3069–3079. [CrossRef]
15. Duan, Z.; Bastiaanssen, W. First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure. *Remote Sens. Environ.* **2013**, *131*, 1–13. [CrossRef]
16. Zhang, Q.; Shi, P.; Singh, V.P.; Fan, K.; Huang, J. Spatial downscaling of TRMM-based precipitation data using vegetative response in Xinjiang, China. *Int. J. Clim.* **2016**, *37*, 3895–3909. [CrossRef]
17. Xu, S.; Wu, C.; Wang, L.; Gonsamo, A.; Shen, Y.; Niu, Z. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. *Remote Sens. Environ.* **2015**, *162*, 119–140. [CrossRef]
18. Chen, C.; Zhao, S.; Duan, Z.; Qin, Z. An Improved Spatial Downscaling Procedure for TRMM 3B43 Precipitation Product Using Geographically Weighted Regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4592–4604. [CrossRef]
19. Lv, A.; Zhou, L. A Rainfall Model Based on a Geographically Weighted Regression Algorithm for Rainfall Estimations over the Arid Qaidam Basin in China. *Remote Sens.* **2016**, *8*, 311. [CrossRef]
20. Chen, Y.; Huang, J.; Sheng, S.; Mansaray, L.R.; Liu, Z.; Wu, H.; Wang, X. A new downscaling-integration framework for high-resolution monthly precipitation estimates: Combining rain gauge observations, satellite-derived precipitation data and geographical ancillary data. *Remote Sens. Environ.* **2018**, *214*, 154–172. [CrossRef]
21. Sharifi, E.; Saghafian, B.; Steinacker, R. Downscaling satellite precipitation estimates with multiple linear regression, artificial neural networks, and spline interpolation techniques. *J. Geophys. Res. Atmos.* **2019**, *124*, 789–805. [CrossRef]
22. Ma, Z.; Xu, J.; He, K.; Han, X.; Ji, Q.; Wang, T.; Xiong, W.; Hong, Y. An updated moving window algorithm for hourly-scale satellite precipitation downscaling: A case study in the Southeast Coast of China. *J. Hydrol.* **2020**, *581*, 124378. [CrossRef]

23. Zhang, Q.; Singh, V.P.; Li, J.; Jiang, F.; Bai, Y. Spatio-temporal variations of precipitation extremes in Xinjiang, China. *J. Hydrol.* **2012**, *434*, 7–18. [CrossRef]
24. Xu, J.; Chen, Y.; Li, W.; Liu, Z.; Tang, J.; Wei, C. Understanding temporal and spatial complexity of precipitation distribution in Xinjiang, China. *Theor. Appl. Climatol.* **2016**, *123*, 321–333. [CrossRef]
25. Tan, X.; Shao, D. Precipitation trends and teleconnections identified using quantile regressions over Xinjiang, China. *Int. J. Clim.* **2016**, *37*, 1510–1525. [CrossRef]
26. 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]
27. Lu, X.; Tang, G.; Wei, M.; Yang, L.; Zhang, Y. Evaluation of multi-satellite precipitation products in Xinjiang, China. *Int. J. Remote Sens.* **2018**, *39*, 7437–7462. [CrossRef]
28. Lu, X.; Wei, M.; Tang, G.; Zhang, Y. Evaluation and correction of the TRMM 3B43V7 and GPM 3IMERGM satellite precipitation products by use of ground-based data over Xinjiang, China. *Environ. Earth Sci.* **2018**, *77*, 209. [CrossRef]
29. Yin, Z.-Y.; Liu, X.; Zhang, X.; Chung, C. Using a geographic information system to improve Special Sensor Microwave Imager precipitation estimates over the Tibetan Plateau. *J. Geophys. Res. Space Phys.* **2004**, *109*, 109. [CrossRef]
30. Yin, Z.-Y.; Zhang, X.; Liu, X.; Colella, M.; Chen, X. An Assessment of the Biases of Satellite Rainfall Estimates over the Tibetan Plateau and Correction Methods Based on Topographic Analysis. *J. Hydrometeorol.* **2008**, *9*, 301–326. [CrossRef]
31. Sørensen, R.; Seibert, J. Effects of DEM resolution on the calculation of topographical indices: TWI and its components. *J. Hydrol.* **2007**, *347*, 79–89. [CrossRef]
32. Eliassem, A. *Provisional Report on Calculation of Spatial Covariance and Autocorrelation of the Pressure Field*; Report No. 5; Videnskaps-Akademiets Institutt for Vaer-Og Klimaforskning: Oslo, Norway, 1954.
33. Gandin, L. *Objective Analysis of Meteorological Fields*; Israel Program for Scientific Translations: Jerusalem, Israel, 1965; 242p.
34. Xie, P.; Xiong, A.-Y. A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. *J. Geophys. Res. Space Phys.* **2011**, *116*, D21106. [CrossRef]
35. Pan, Y.; Shen, Y.; Yu, J.; Zhao, P. Analysis of the combined gauge-satellite hourly precipitation over China based on the OI technique. *Acta Meteorol. Sin.* **2012**, *70*, 1381–1389. (In Chinese)
36. Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [CrossRef]
37. Propastin, P.; Kappas, M.; Erasmi, S. Application of geographically weighted regression to investigate the impact of scale on prediction uncertainty by modelling relationship between vegetation and climate. *Int. J. Spat. Data Infra. Res.* **2008**, *3*, 73–94.
38. Lu, X.; Tang, G.; Wang, X.; Liu, Y.; Jia, L.; Xie, G.; Li, S.; Zhang, Y. Correcting GPM IMERG precipitation data over the Tianshan Mountains in China. *J. Hydrol.* **2019**, *575*, 1239–1252. [CrossRef]
39. Ma, Z.; Shi, Z.; Zhou, Y.; Xu, J.; Yu, W.; Yang, Y. A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed. *Remote Sens. Environ.* **2017**, *200*, 378–395. [CrossRef]

Article

Assessing the Impact of GNSS ZTD Data Assimilation into the WRF Modeling System during High-Impact Rainfall Events over Greece

Christos Giannaros [1],*, Vassiliki Kotroni [1], Konstantinos Lagouvardos [1], Theodore M. Giannaros [1] and Christos Pikridas [2]

[1] National Observatory of Athens, Institute of Environmental Research and Sustainable Development, 15236 Athens, Greece; kotroni@noa.gr (V.K.); lagouvar@noa.gr (K.L.); thgian@noa.gr (T.M.G.)
[2] Department of Geodesy and Surveying, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; cpik@topo.auth.gr
*   Correspondence: chrisgiannaros@noa.gr;

Abstract: The derivation of global navigation satellite systems (GNSSs) tropospheric products is nowadays a state-of-the-art technique that serves both research and operational needs in a broad range of applications in meteorology. In particular, GNSS zenith tropospheric delay (ZTD) data assimilation is widely applied in Europe to enhance numerical weather predictions (NWPs). The current study presents the first attempt at introducing assimilation of ZTDs, derived from more than 48 stations of the Hellenic GNSS network, into the operational NWP system of the National Observatory of Athens (NOA) in Greece, which is based on the mesoscale Weather Research and Forecasting (WRF) model. WRF was applied during seven high-impact precipitation events covering the dry and wet season of 2018. The simulation employing the ZTD data assimilation reproduces more accurately, compared to the control experiment, the observed heavy rainfall (especially for high precipitation events, exceeding 20 mm in 24h) during both dry and wet periods. Assimilating ZTDs also improves the simulation of intense (>20 mm) convective precipitation during the time window of its occurrence in the dry season, and provides a beneficial influence during synoptic-scale events in the wet period. The above results, which are statistically significant, highlight an important positive impact of ZTD assimilation on the model's precipitation forecast skill over Greece. Overall, the modelling system's configuration, including the assimilation of ZTD observations, satisfactorily captures the spatial and temporal distribution of the observed rainfall and can therefore be used as the basis for examining further improvements in the future.

Keywords: data assimilation; WRF model; high-impact rainfall events; GNSS ZTD; assessment

1. Introduction

The use of global navigation satellite systems (GNSSs) is essential in a variety of fields that require precise location and time information, including aviation (e.g., Sabatini et al. [1]), transportation (e.g, Kubo et al. [2]), search and rescue services (e.g, Molina et al. [3]), agriculture (e.g., Kahveci et al. [4]), and maritime operations (e.g., Ostolaza et al. [5]). Additionally, remote sensing of atmospheric constituents with the exploitation of GNSS signals is nowadays a well-established and widely applied approach, which is referred to as GNSS meteorology [6]. The methodology is based on the fact that the radio signals transmitted from the satellites to the receivers on the ground are delayed when propagating through the troposphere due to the presence of dry gases and water vapor [7]. Advanced GNSS processing techniques produce various tropospheric products that are used in several meteorological applications, including nowcasting and numerical weather prediction (NWP) [8–11],

as well as weather monitoring, including extreme events [12–14]. Special meteorological interest derives from the near-real time (NRT) ZTDs, which are estimated based on raw GNSS observations. The ZTD is a standard GNSS product expressing the total signal delay in the zenith direction above a receiver [6,15]. This vertical lag contains information on the total columnar amount of water vapor [16].

In Europe, collaborative scientific efforts over the past two decades substantially contributed to the development of networks and analysis centers collecting and processing, respectively, GNSS data to compute tropospheric delays. The establishment of the European GNSS water vapor program (E-GVAP; [17]) in 2005 allowed for the operational distribution of NRT ZTD estimates to the meteorological community [6]. This service encouraged the implementation of precipitation forecast impact studies involving the assimilation of NRT ZTD observations into NWP models. Poli et al. [18] found an improvement in the prediction of precipitation patterns over France during spring and summer when assimilating ZTD data into a global four-dimensional variational (4D-var) assimilation and forecasting system. Positive impacts of ZTD data assimilation on precipitation forecasts of convective-scale systems over France were also demonstrated by Yan et al. [19,20], using the three-dimensional variational (3D-var) assimilation systems of AROME, ALADIN, and Meso-NH (Mesoscale Non-Hydrostatic) atmospheric models. Eresmaa et al. [21] positively evaluated the application of a bias correction algorithm in the ZTD observations prior to assimilating them into the 3D-var/HIRLAM (HIgh Resolution Limited Area Model) modeling system. In addition, their study showed mixed results (positive/negative) concerning the impact of ZTD assimilation on the predicted precipitation, which were dependent on the forecast lead-time. Similarly, Bennitt and Jupp [22], and Schwitalla et al. [23], who performed numerical experiments over Europe with the United Kingdom (UK) Meteorological Office 3D and 4D-var model, and the 3D-var/WRF model, respectively, found no clear impact of ZTD assimilation on precipitation forecasts. Bennitt and Jupp [22] further showed that 4D-var assimilation does not lead to better forecasts compared to the assimilation under the 3D-var system. Boniface et al. [24] identified that the impact of ZTD assimilation in AROME model depends on the rainfall synoptic conditions, while Arriola et al. [16] highlighted the importance of ZTD observations processing before the assimilation application. More recent studies focused on investigating different observational bias correction methods and on estimating spatiotemporal correlations of observations errors for application on ZTD data used for data assimilation [25–27].

The literature review shows that improvements in the precipitation forecast skill may be gained by assimilating ZTD observations into NWP models. However, regional studies using the WRF modeling system are relatively rare (e.g., Rohm et al. [28]), while no research focusing on Greece has ever been conducted. The present study aims to fill this knowledge gap by examining the impact of GNSS ZTD data assimilation on the precipitation forecast skill of the 3D-var/WRF model. For that reason, four dry and three wet period heavy precipitation events, observed in 2018 over Greece, were selected to examine the model's performance under different meteorological conditions. Prior to assimilation, an integrated method was employed to process the ZTD observations in order to review their quality, correct their systematic errors with respect to the model, and finally, to use those that fulfilled specific selection criteria. The results of the numerical experiments are compared against precipitation observations obtained from the network of surface meteorological stations operated by NOA [29]. The applied model configuration is based on NOA's operational weather forecasting system in order to assess the GNSS ZTD data assimilation capacity in terms of supporting real-time weather prediction applications.

2. Materials and Methods

2.1. Case Studies

The annual distribution of precipitation over Greece is characterized by significant spatial and temporal variations. Following seasonal variability, a hydrological year in Greece can be divided into a dry (summer) and wet (autumn, winter and spring) period with much greater precipitation

amounts and resulting impacts occurring during the latter [30,31]. Rainfall during the dry period of the year is primarily attributed to atmospheric instability, leading to convective storms [32,33]. Precipitation in autumn and winter is mainly related to intense cyclone activity [34,35], but also to less intense cyclones with relatively long lasting embedded meso-scale convective systems that interact with the complex topography [36]. Considering the spatiotemporal precipitation pattern described above, four dry and three wet period rainfall events observed in 2018 were selected for the purposes of the current study. The precipitation intensity was also taken into account by selecting high-impact episodes, as summarized in Table 1.

Table 1. Synopsis of the 2018 selected case studies and their associated driving meteorological mechanisms and impacts.

Period	Case Study	Driving Meteorological Mechanism	Maximum Total Rainfall	Impacts
Dry	May 10	Unstable atmospheric conditions resulting in strong convective activity	74 mm in 1 h over Thessaloniki	Flash flooding; houses, monuments, electricity supply and transportation were affected
	June 16	Unstable atmospheric conditions resulting in strong convective activity	~ 65 mm in 1 h over Skopelos island	Major damages in housing, crops, rural roads and electricity supply
	July 26	Unstable atmospheric conditions resulting in strong convective activity	> 100 mm in less than 2 h over northern suburbs of Athens	Flash flooding; houses, vehicles and transportation were affected
	August 27–28	Unstable atmospheric conditions resulting in strong convective activity.	> 150 mm in 24 h over northern Greece and > 90 mm in 24 h over central Greece and Crete	Flash flooding; houses, telecommunications, electricity supply and transportation were affected
Wet	September 29–30	Low-pressure system with tropical characteristics (Medicane; [37,38])	~ 500 mm in 48 h over Phiotis region	Floods and landslides; four dead people; extensive damages in buildings and public infrastructures
	November 18	Low-pressure system	133 mm in 24 h over northern Greece	Flash flooding; crops and road network were affected
	November 26–27	Low-pressure system	161 mm in 48 h over northwest Greece	Floods and landslides; housing, public infrastructures and road network were affected

2.2. The WRF NWP System

The modeling system used in the current work is WRF-ARW (advanced research WRF), version 4.0.2 [39], which is the core of the integrated HERMES modeling system that operates daily at NOA to support operational weather forecasting and natural disaster early warnings [40,41].

The WRF setup was based on two 2-way nested modeling domains, as shown in Figure 1a. The coarse domain (d01) was used to simulate the synoptic-scale atmospheric conditions with a spatial resolution of 10 km (mesh size of 500×500), while the innermost domain (d02) focused on the study area (Greece,) having a horizontal grid resolution of 2 km (mesh size of 551×551). The vertical structure of both domains included 41 unevenly spaced hybrid sigma-pressure layers up to 50 hPa. Initial and boundary conditions were obtained from the Global Forecast System (GFS) operational surface and upper air atmospheric analysis data at 0.25×0.25 spatial and 6 h temporal resolution, provided by the National Center for Environmental Prediction (NCEP). The selected physics parameterizations are presented in Table 2.
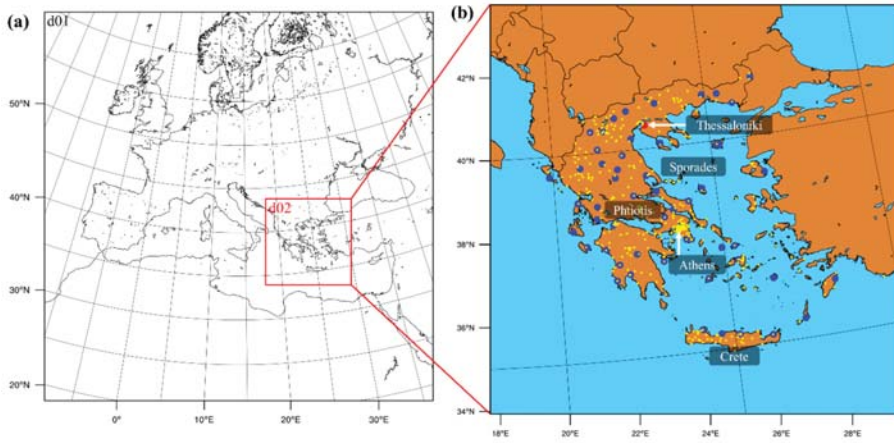
Figure 1. (**a**) The applied WRF domains' configuration and (**b**) study domain (d02) with the locations of the GNSS stations (blue dots; the red dot highlights the AUT1 station used for examining the NRT ZTDs' accuracy) and ground-based meteorological sites (yellow dots) utilized for the November 27 precipitation event, and with identification of regions where high rainfall amounts were observed during the examined episodes.

Table 2. Summary of the applied WRF model physics.

Physics	Parameterization	References
Microphysics	Thompson	Thompson et al. [42]
Convection [1]	Kain-Fritsch	Kain [43]
Planetary boundary layer	Yonsei University	Hong et al. [44]
Surface layer	Revised MM5	Jiménez et al. [45]
Land surface	Noah	Tewari et al. [46]
Short- and long-wave radiation	RRTMG	Iacono et al. [47]

[1] Convection parameterization was used only for domain d01.

2.2.1. Data Assimilation Scheme

The WRF data assimilation package provides various assimilation techniques, ranging from empirical (e.g., nudging) and statistical (e.g., variational analysis) approaches, to advanced methods (e.g., ensemble Kalman filter). In the variational scheme, the best possible estimate of the model's initial state (analysis) is determined by the minimization of a prescribed cost function, $J(x)$, that combines a background forecast (first-guess), observations and estimates of both modeling and observational error covariances. 3D and 4D-var options are available in WRF [48,49]. The main difference between the two implementations is that the observations in 4D-var/WRF model are integrated within an assimilation window at the exact time of the observations. Consequently, greater computing resources compared to 3D-var assimilation are necessary [50]. Taking the possible application of GNSS ZTD data assimilation in NOA's operational forecasting system into account, without significantly altering the timeliness of forecast delivery, the 3D-var option was employed in the present work.

According to Barker et al. [48], the cost function in WRF 3D-var operation can be written as:

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(H(x) - y^o)^T R^{-1}(H(x) - y^o)$$

expressing the weighted distance of analysis, x, to first-guess, x_b, and observations, y^o. The contributions of x_b and y^o to x are determined by the model background errors covariance matrix, B, and the observations errors covariance matrix, R, respectively. The observations operator H is used to transform the model's analysis to observational space. The WRF first guess can be either the initial analysis computed based on global fields (cold start) or a previous model forecast (cycling). In the last-mentioned case, the observations are used to improve the current model state, ultimately resulting in better analysis and prediction. The 3D-var/WRF system supports the assimilation of in-situ conventional measurements, remotely sensed observations, and satellite radiances. Ground-based GNSS ZTD data are included in the remote sensing observations category and they are handled by the GPSPW operator [28,39].

The success of the 3D-var assimilation depends heavily on the accuracy with which the observations and model background errors are specified. The R matrix is determined by instrument and representation errors. The latter includes errors introduced by the observations pre-processing and operator, as well as by the effect of unresolved scales in the model [51]. The observations preprocessor (OBSPROC) in the 3D-var/WRF system defines the R matrix, based on pre-specified observations errors. These error values had to be determined by the user for the case of GNSS ZTD data. No correlation in space and time between the individual observation errors is assumed by the OBSPROC package [39]. In the current study, the GNSS ZTD formal errors, derived during the GNSS raw observations processing, were used in the 3D-var/WRF application. These errors are related to uncertainties induced by satellite orbits, antennas, signals multipath, ionospheric delays, and the applied mapping function. Hence, they are a measure of the ZTD estimation uncertainty that considers both measurement and processing errors. The B matrix is considered static and its specification is based on "climatological" estimates, assuming that the model background errors correlations are homogeneous and isotropic. Its computation involves the application of the National Meteorological Center (NMC) method [52], in which the covariances for a set of five independent control variables are derived for a given domain by averaging the forecast differences between the 24 and 12 h prediction lead time over a period of time [39,48]. In the present work, the B matrix was calculated for each rainfall event, considering that the model background fields are different for each case study. In particular, the B matrix was calculated based on a series of WRF simulations of 24 h duration for the domain of interest (d02), which were initialized twice daily at 0000 UTC and 1200 UTC and covered a 10 day period, starting thus 10 days before each episode occurrence. Similar approaches for computing the B matrix have also been applied in previous studies [28,53].

2.2.2. Numerical Experiments

Figure 2 illustrates the flow diagrams of the numerical experiments implemented in the current study. Two simulations were performed for each day of the examined events. The first one was carried out without assimilation and refers to the control (CTL) experiment. The second simulation refers to the GNSS ZTD observations assimilation (ZTD) experiment, in which the 3D-var/WRF system under cycling mode was employed. The criterion for selecting the time of ZTD data assimilation for the dry period events was mainly based on the approximate time of the actual start of the convective precipitation and it was set 6 h prior to this time. Thus, data assimilation was applied at 0600 UTC for the May and July cases (Figure 2b), and at 0000 UTC for the June and August cases. For the wet season episodes, the 0000 UTC was chosen as the assimilation hour because the surface low-pressure systems were well developed at that time (Figure 1a).
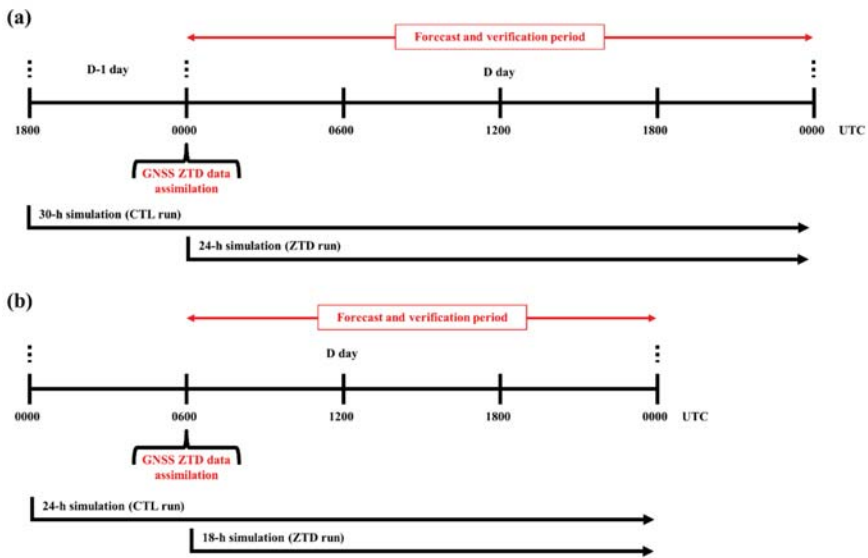
Figure 2. Flow diagrams illustrating the performed WRF simulations without and with ZTD data assimilation for (**a**) 24 h and (**b**) 18 h forecast and verification periods.

2.3. ZTD Observations

GNSS applications in meteorology require careful characterization and assessment of tropospheric products derived from ground-based GNSS receivers. This is one of the key specialization areas for the GNSS Quality Control (GNSS-QC) team of the Aristotle University of Thessaloniki (AUTh), which processes and monitors a network of permanent GNSS stations located in Greece. The network consists of over 90 GNSS reference stations deployed by both public and private organizations. Four out of the 90 GNSS stations are included in the EUREF (European reference frame; [54]) permanent network (EPN). The GNSS-QC group has contributed to the E-GVAP project since 2014, providing NRT ZTD estimates to the European NWP centers [55–57].

In the current study, hourly ZTD observations, at 0000 UTC and 0600 UTC, of the AUTh GNSS-QC network were used for the WRF data assimilation experiments. The data collection and analysis were carried out for 2018 in the framework of BERTISS (Balkan-Mediterranean Real Time Severe Weather Service) project. An automated NRT processing scheme was employed using the Bernese software package, version 5.2 [58], and several coding scripts and algorithms for computations, data management, quality control, and database upload [15,59].

Since the observations accuracy and their associated errors is crucial in improving the model analysis, the quality control of the GNSS ZTD estimates and the appropriate treatment of ZTD formal errors within the 3D-var/WRF system are necessary. For this, a thorough observations pre-processing was applied in the present work to prepare the ZTD data for the assimilation process.

Data Pre-Processing

The accuracy of the NRT ZTDs, derived from the AUTh GNSS network, was examined by comparing the data against high-quality reference ZTD data provided by EPN. Based on the availability of the EPN weekly combined ZTD solutions, the validation was performed over the AUT1 station (Figure 1b) using concurrent data of the 2018 dataset. Following previous studies [60–62], the accuracy of the ZTD observations was defined as the standard deviation of the difference between the AUT1 and EPN ZTDs. The analysis resulted in an accuracy of 7.6 mm, which is between the optimal (5 mm) and

target (10 mm) accuracy value according to the E-GVAP standards for NRT ZTDs used in the NWP models [63]. This outcome was a strong indication that the accuracy of the AUTh GNSS ZTD data was sufficient to assimilate them into the WRF modeling system.

Further, the observations in the 3D-var system are considered to have unbiased errors with respect to the WRF model. To meet this assumption, a statistical bias correction was applied to the ZTD data for each studied precipitation episode. More specifically, the differences between the observed and modeled ZTDs were computed for each GNSS receiver during a 10 day period prior to each event using the WRF simulations conducted for the model background errors covariance matrix specification (see Section 2.2.1). Then, the corrections were estimated as the mean values of the ZTD differences and they were subtracted from the ZTD observations that were lined up for data assimilation. Even though this method provides statistical corrections, this is a standard bias correction approach for ZTD data that proved to be successful in reducing the observations-model divergences and capturing the systematic errors between the ZTD observations and model forecasts [16,18–25,28]. The last stage of the ZTD data pre-processing included a selection algorithm based on the following conditions: (a) The formal ZTD error to be lower than the standard deviation of the difference between the observed and modeled ZTD, (b) the ZTD difference between observations and model output to be lower than five times the ZTD formal error, and (c) the difference between the receiver height and the model's orography to be below 100 m. Similar criteria have been applied in previous studies [19–25]. Based on data availability and the above selection algorithm, the number of GNSS reference stations used per event ranged from 48 to 56. The assimilated ZTD observations had the same range, since assimilation was performed once (at 0000 UTC or 0600 UTC) during the conducted experiments. The locations of the 55 stations used for the event of 27 November 2019 are shown in Figure 1b, demonstrating a sufficient and homogeneous spatial coverage of the Greek territory.

2.4. Evaluation Process

Hourly precipitation measurements covering the entire study area were used to evaluate the model performance under both the CTL and ZTD numerical experiments. The data were collected from a dense network of weather stations operated by NOA [29]. In terms of data availability, different rain gauges were used for each case study. In total, 340 to 360 rain gauges were utilized for each rainfall event (Figure 1b). The observed and modeled precipitation data were paired in time and space, considering the nine nearest to the location of each rain gauge model grid points. The grid point having the closest predicted value to the observed one was selected for evaluation in order to avoid penalizing the model performance due to possible small spatial displacements of rainfall. The evaluation was performed for the 24 h and 6 h accumulated precipitation at 0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC.

Using the observation-model pairs, qualitative statistical measures were computed on the basis of a dichotomous application system (occurrence/no occurrence of precipitation) for six distinct rain thresholds: above 0.2, 1, 2, 5, 10, and 20 mm. The computed scores included the probability of detection (POD), the false alarm ratio (FAR), the equitable threat score (ETS), and the frequency bias (FBIAS). POD shows the fraction of observed events that were correctly modeled and ranges from zero (0: wrong forecast) to one (1: perfect forecast). FAR is the fraction of forecast events that were not observed and spans from zero (0: perfect forecast) to one (1: wrong forecast). ETS measures the skill of a model prediction considering the chance of randomly correct forecasts. ETS values close to unity (1) indicate a high-accuracy forecast, whereas ETS values that are close to zero (0), or even negative, shows a poor or random forecast quality. FBIAS is the ratio between the frequencies of forecasted and observed events and indicates whether a model tends to underestimate (FBIAS < 1) or overestimate (FBIAS > 1) the frequency of the occurrence of the observed events. To determine the statistical significance of the qualitative score differences between the conducted experiments, a hypothesis test approach was applied, using two confidence intervals: (i) 90% and (ii) 95%. The test was based on the construction of a probability density function that was consistent with the assumption that there was no difference

between the qualitative statistical measures computed using the CTL simulations and those calculated based on the ZTD simulations. A brief description of the implemented method can be found in Giannaros et al. [64]. Quantitative statistical measures were also calculated, namely the mean bias (MB) and mean absolute error (MAE), for each precipitation threshold in order to account for the magnitude of errors. MB is used as a measure of the model tendency of rain underestimation (MB > 0) or overestimation (MB < 0), while MAE represents the absolute deviation between the observational and modeled precipitation. Following Giannaros et al. [64], the statistical significance of the MAE differences between the CTL and ZTD experiments was computed by applying the non-parametric Wilcoxon signed-rank test at the 90% and 95% confidence intervals. In addition to the statistical evaluation, the modeled differences concerning the rainfall distribution were thoroughly investigated for two events representing different synoptic conditions.

3. Results

3.1. Statistical Evaluation

3.1.1. Daily Precipitation

Figure 3 presents the categorical scores for each 24 h accumulated precipitation threshold, aggregated for the dry and wet season rainfall events for the CTL and ZTD numerical experiments. Overall, the model performs adequately in capturing the occurrence (non-occurrence) of precipitation, as indicated by the POD (FAR), which is higher (lower) than 0.57 (0.38) and 0.67 (0.26) for all thresholds during the dry (Figure 3a) and wet (Figure 3b) period, respectively. The ETS values show a satisfying precipitation forecast quality, especially for the wet season (Figure 3b), when they range from 0.53 to 0.88. During both periods, the FBIAS values demonstrate that the model underestimates the observed frequency of higher than 20 mm daily precipitation, whereas it slightly overestimates the frequency of the observed daily rainfall for the lower than 20 mm thresholds, except those that are greater than 2 mm (5 mm and 10 mm) during the CTL experiment in the wet (dry) season (Figure 3).

Assimilating ZTD observations into the 3D-var/WRF model leads to increased probability of precipitation detection during the dry period (Figure 3a) across all rainfall thresholds. Especially for the highest precipitation threshold (>20 mm), the ZTD assimilation induced relative improvement is 10.6% (statistically significant at the 95% confidence interval; Figure 3a). Concerning FBIAS, the ZTD experiment leads to larger frequency biases compared to the CTL simulation, when precipitation is lower than 20 mm in the dry period, reaching 10.3% relative difference for the above 10 mm rainfall threshold (statistically significant at the 90% confidence interval; Figure 3a). FAR in the dry period is also higher for the same threshold (>10 mm) during the ZTD experiment, whereas a decrease in FAR is evident for greater than 20 mm 24 h precipitation when ZTD assimilation is applied (Figure 3a). No marked differences between the CTL and ZTD experiments are evident during the dry season for ETS, except from the considerable improvement by 11.4% (statistically significant at the 90% confidence interval) provided by the ZTD assimilation for the highest precipitation threshold (Figure 3a). For the same threshold, the ZTD experiment also leads to statistically significant improvements for all statistical measures, except FBIAS during the wet season (Figure 3b). In particular, POD and ETS are increased by 3.5% and 8.1% at the 95% confidence level, while FAR is reduced by 16.9% at the 90% confidence interval (Figure 3b).

(a)

(b)
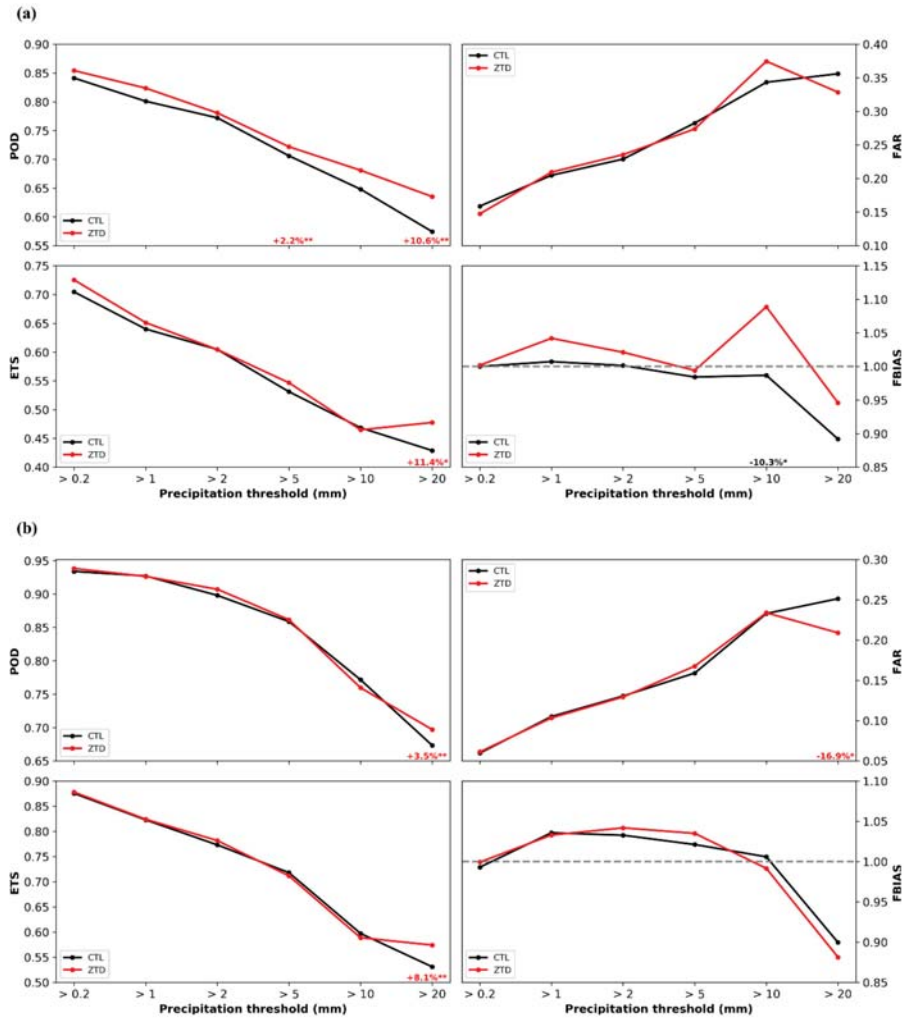
Figure 3. Qualitative model performance statistics averaged for the (**a**) dry and (**b**) wet period events for daily precipitation under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).

In terms of quantitative statistics, the model overestimates, to a small extent, the 24 h accumulated precipitation values that are lower than 10 mm (20 mm) during the dry (wet) period, whereas it significantly underestimates the high rainfall accumulations for both dry and wet season events (Figure 4). More specifically, the CTL and ZTD mean absolute errors exceed 15 mm for the larger precipitation threshold, showing that both numerical experiments cannot capture the magnitude of severe rainfall (Figure 4). A similar extent of errors for intense precipitation thresholds have also been found in previous studies (e.g., see References [53,65,66]), showing that quantitative precipitation forecasting (QPF) remains a challenge for regional NWP systems due to uncertainties associated with physics parameterizations, primarily microphysics and convection, domain configuration (e.g.,

resolution and size), and initial conditions [67,68]. The improvement of a model's initial state through data assimilation results in more accurate QPF. This is evident in the present study, as the ZTD simulations reduce the deviations from the observations in the precipitation interval [20,...) mm by ~1.10 mm during both dry and wet periods (Figure 4). In the latter season, this reduction corresponds to a 5.5% (statistically significant at the 95% confidence interval) relative improvement in MAE (Figure 4b). For the rest of the rainfall thresholds, lower (higher) ZTD MAEs can be seen between 2 (0 mm) and 20 mm (2 mm) of precipitation during the wet (dry) period (Figure 4).
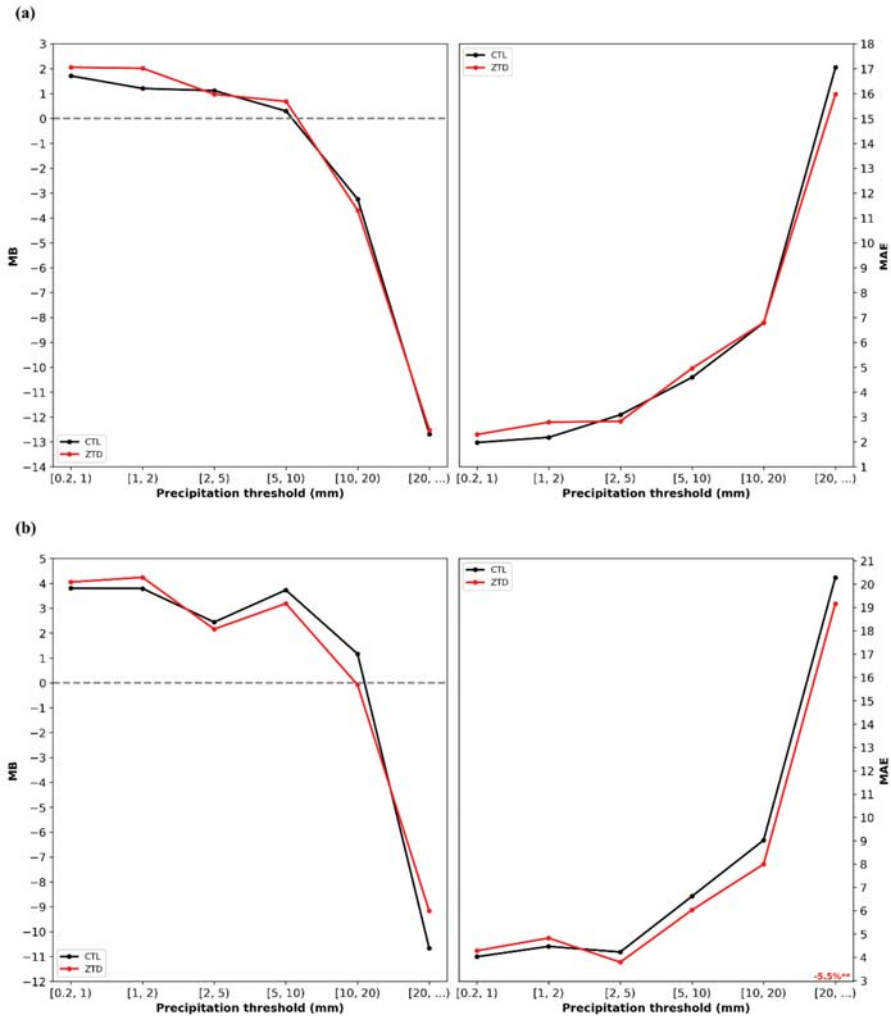
Figure 4. Quantitative model performance statistics averaged for the (**a**) dry and (**b**) wet period events for daily precipitation under six rainfall intervals during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).

3.1.2. 6 h Accumulated Precipitation

Further, the 6 h accumulated precipitation forecasts have been also verified in order to assess the impact of the ZTD data assimilation depending on the forecast lead time (Figures 5–10). The statistical measures for the dry season were computed for the period between 0600 UTC and 1800 UTC. The periods from 0000-0600 UTC and 1800-0000 UTC were discarded from the analysis because less precipitation was observed and fewer observation-model pairs were available compared to the 0600–1800 UTC time window. Figures 5 and 6 illustrate that the ZTD assimilation leads to a marked improvement of the precipitation forecasting in the dry season by increasing the probability of detection and the prediction quality. The improvement is more profound during the afternoon hours (12000–18000 UTC), when statistically significant increases at the 95% confidence level are evident for POD and ETS, which reach 27.8% for higher than 10 mm threshold and 21.4% for higher than 5 mm threshold, respectively (Figure 5b). In the same 6 h interval, the ZTD experiment results in higher FBIAS values, which are greater than 1 when rainfall is lower than 20 mm. This finding partially explains the overall overestimation of the observed events frequency found for the below 20 mm 24 h precipitation in the dry period (Figure 3a). FAR is mainly decreased during the ZTD experiment between 0600 and 1800 UTC in the dry period, especially for the higher rainfall threshold (Figure 5). This is also true for MAE, as shown in Figure 6. In particular, statistically significant reductions of 13% (19.9%) at the 95% (90%) confidence level are introduced by the ZTD simulation for precipitation above 20 mm (between 5 and 10 mm) from 1200–1800 (0600–1200) UTC (Figure 6). Figure 6 also illustrates that, when considering the lowest three rainfall thresholds (<5 mm), the WRF model mainly overestimates the observed precipitation, whereas it underestimates the higher than 5 mm observed rainfall during the examined forecast lead times. From 1200-1800 UTC, the model overestimation between 1 and 2 mm of precipitation is higher by 57.4% (statistically significant at the 90% confidence interval) during the ZTD simulation (Figure 6). The forecast errors for the rainfall thresholds up to 10 mm are smaller than ~9 mm for all 6 h intervals, while in contrast, they are greater than 19 mm for the highest precipitation threshold (Figure 6).
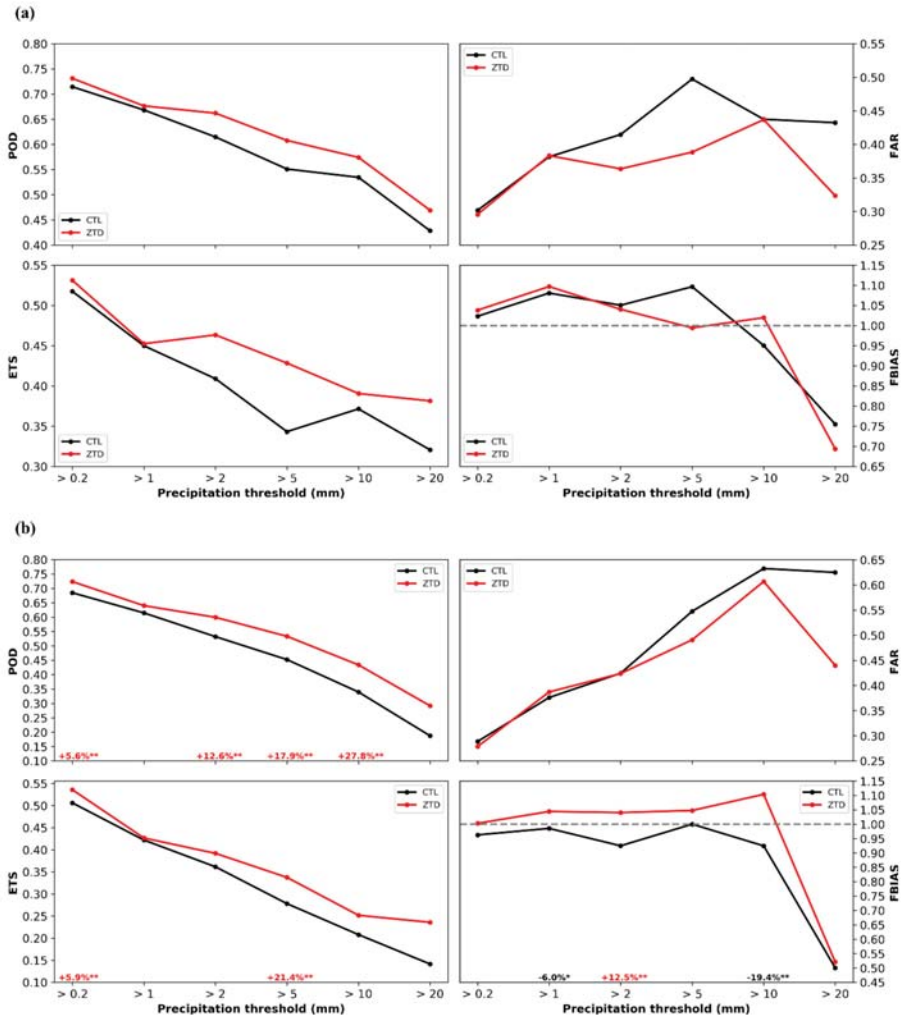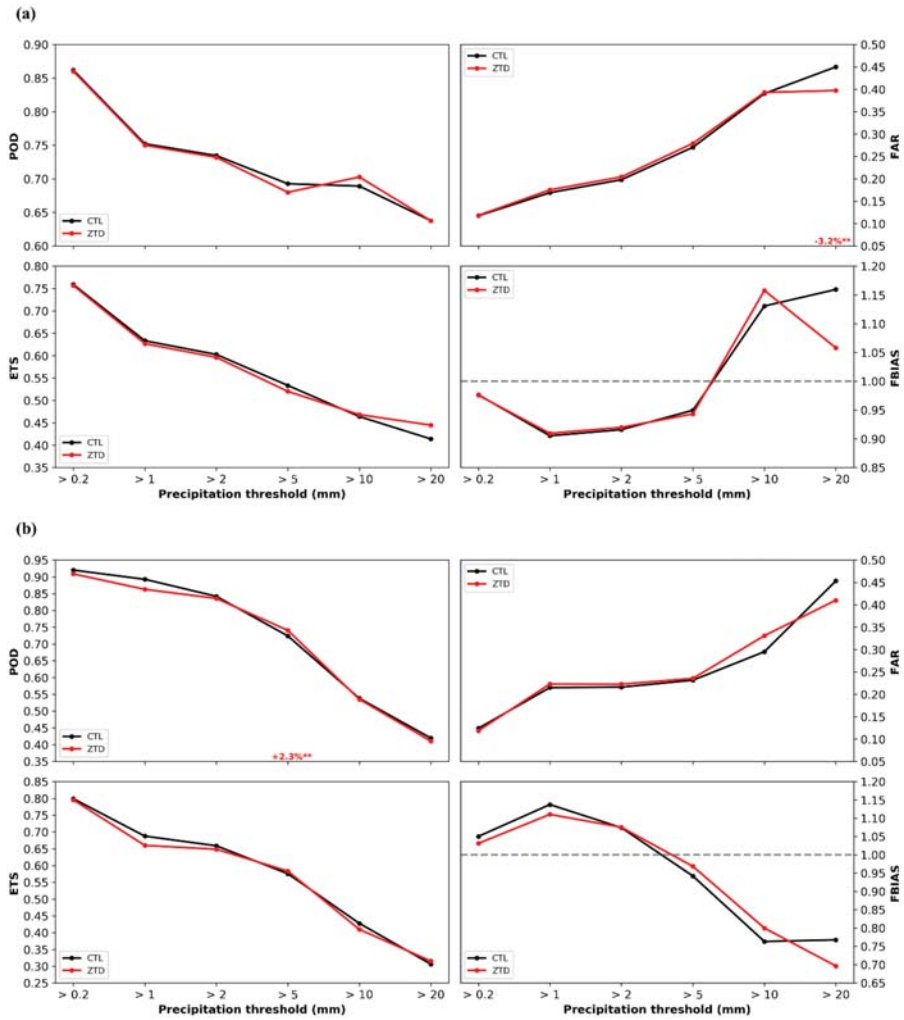
Figure 5. Qualitative model performance statistics averaged for the dry period events for precipitation accumulations between (**a**) 0600–1200 UTC and (**b**) 1200–1800 UTC under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).
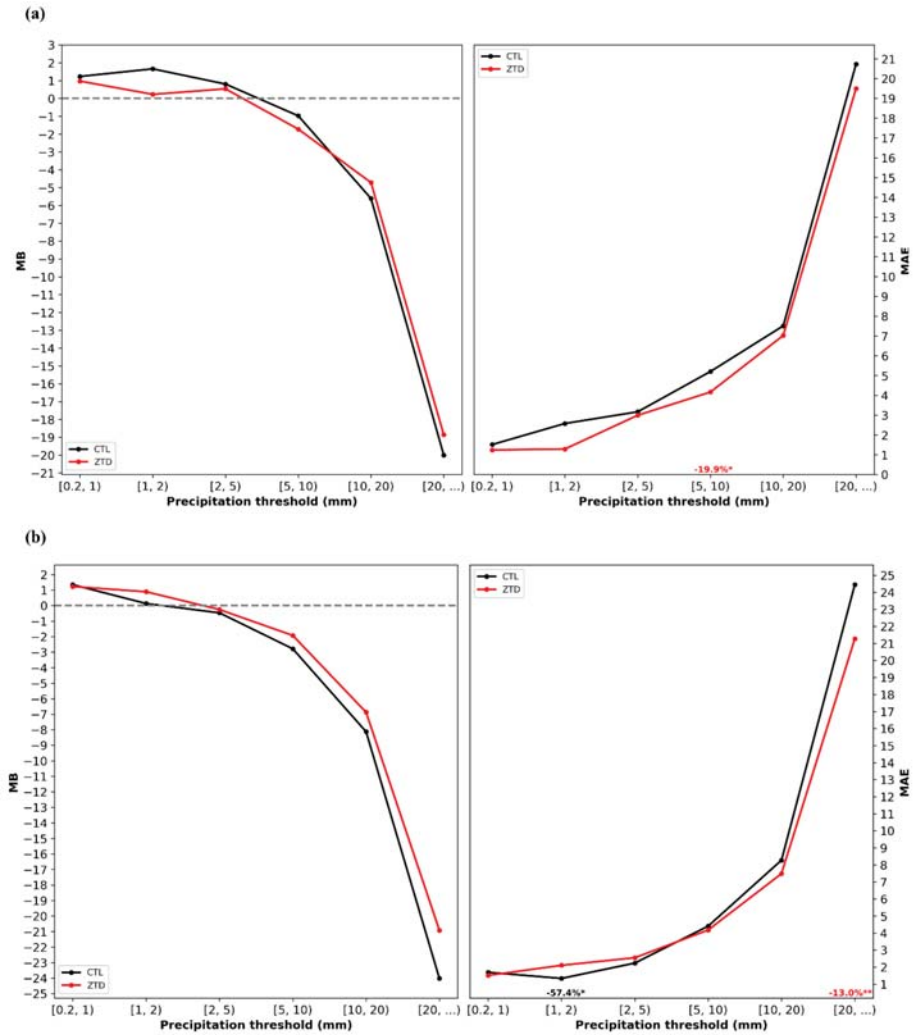
Figure 6. Quantitative model performance statistics averaged for the dry period events for precipitation accumulations between (**a**) 0600-1200 UTC and (**b**) 1200-1800 UTC under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).
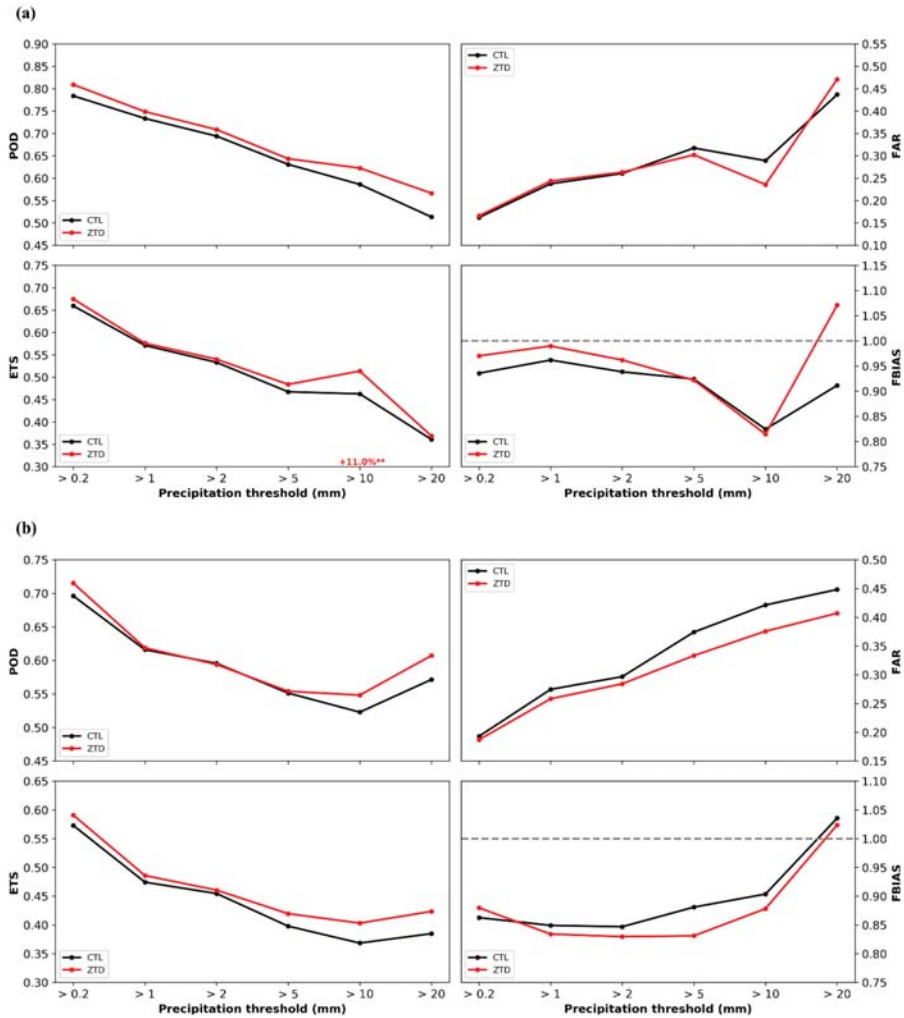
Figure 7. Qualitative model performance statistics averaged for the wet period events for precipitation accumulations between (**a**) 0000–0600 UTC and (**b**) 0600–1200 UTC under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).

Figure 8. Qualitative model performance statistics averaged for the wet period events for precipitation accumulations between (**a**) 1200–1800 UTC and (**b**) 1800–0000 UTC under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).
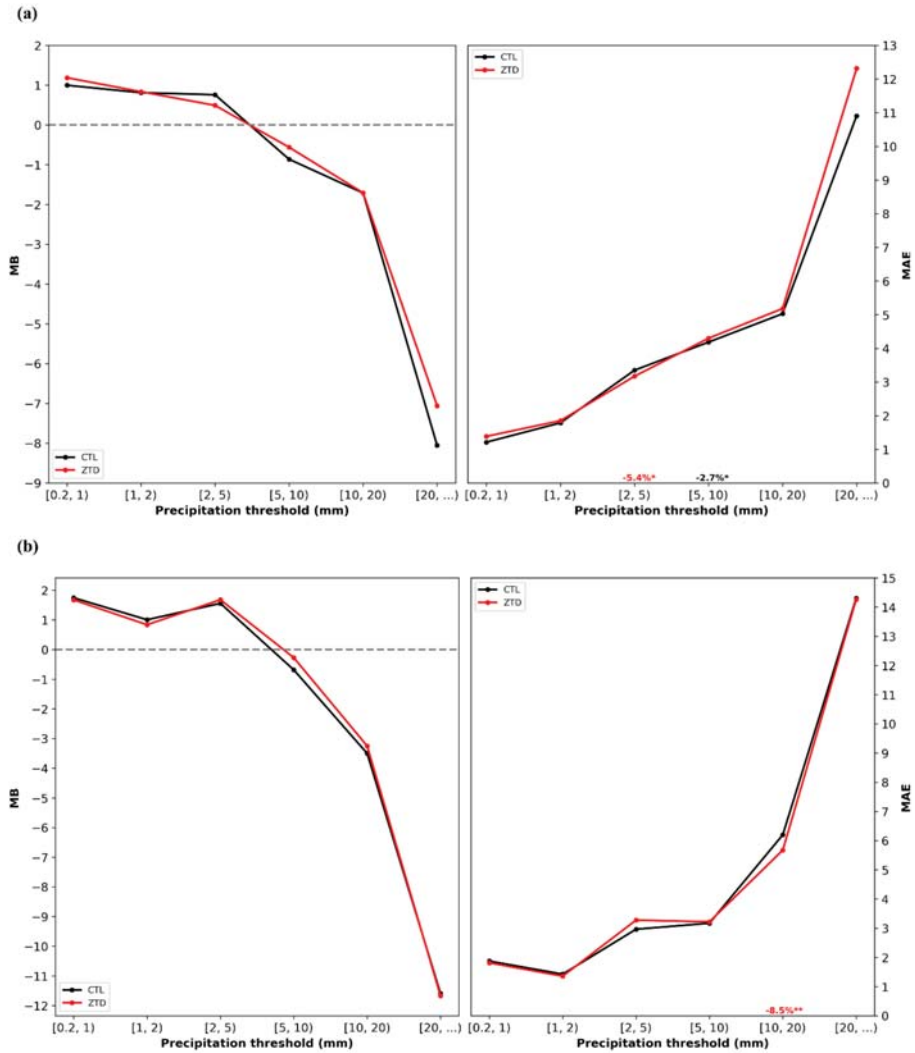
Figure 9. Quantitative model performance statistics averaged for the wet period events for precipitation accumulations between (**a**) 0000–0600 UTC and (**b**) 0600–1200 UTC under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).
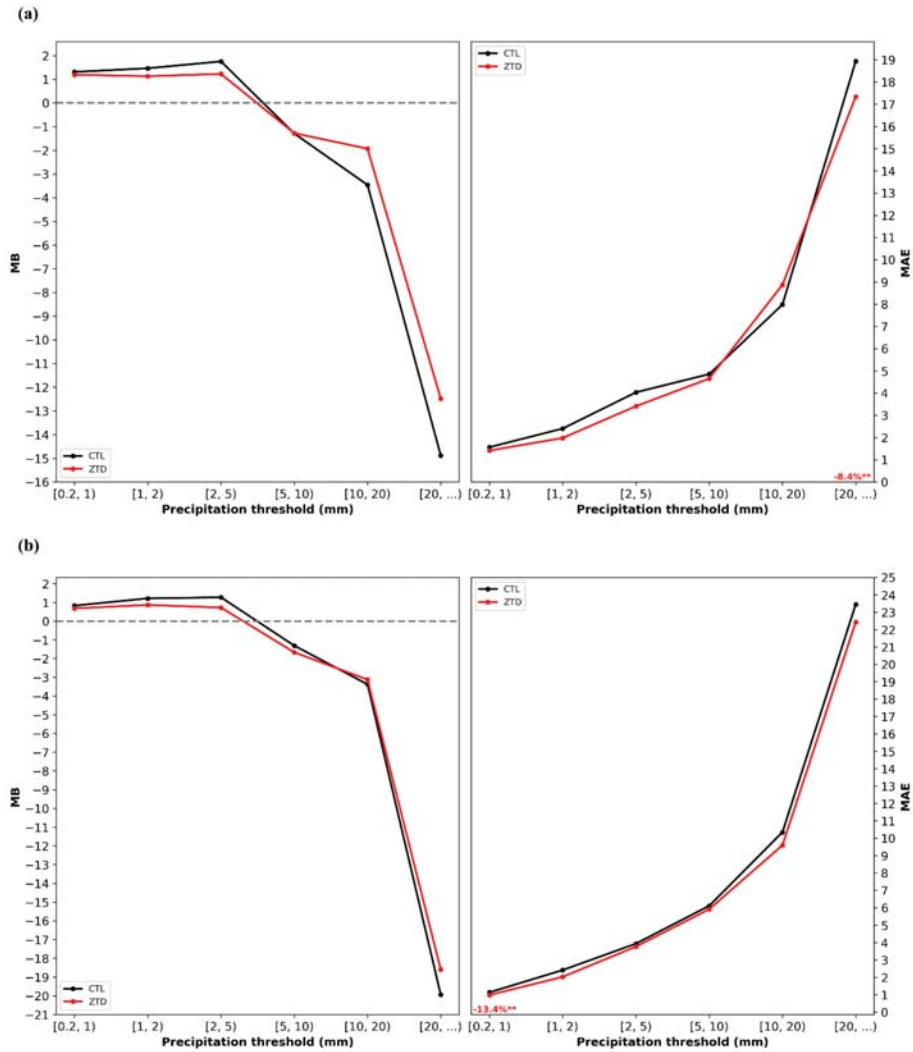
Figure 10. Quantitative model performance statistics averaged for the wet period events for precipitation accumulations between (**a**) 1200–1800 UTC and (**b**) 1800–0000 UTC under six rainfall thresholds during the CTL and ZTD numerical experiments. Percentages indicate the relative difference of the statistical measures between the conducted experiments (one asterisk shows statistical significance at the 90% confidence interval, while two asterisks show statistical significance at the 95% confidence interval).

During the wet season (Figures 7–10), the positive impact of ZTD assimilation on the 6 h accumulated precipitation, especially when exceeding 20 mm, is clearly shown. More specifically, during the first 6 h of the numerical forecasts, the FAR is decreased by 3.2% (statistically significant at the 95% confidence interval) for the highest rainfall threshold when ZTD data are assimilated in the WRF model (Figure 7a). For the same period and threshold, FBIAS is closer to 1 during the ZTD experiment, whereas no significant divergences between the conducted simulations are found for POD and ETS (Figure 7a). Marked differences are also not evident between 0600 UTC and 1200 UTC for all qualitative statistical measures and precipitation thresholds, except for POD, which is higher by 2.3%

(statistically significant at the 95% confidence interval) during the ZTD experiment, when rainfall is above 5 mm (Figure 5b). From 1200-000 UTC, the POD and ETS scores are higher for the ZTD compared to the CTL experiment for the majority of the precipitation thresholds. Especially for the greater rainfall threshold, the improvement provided by the ZTD assimilation in POD and ETS is 10% in the intervals 1200-1800 UTC and 1800-0000 UTC, respectively (Figure 8). Additionally, a statistically significant (95% confidence level) reduction of ETS by 11% is evident during the ZTD simulation for the higher than 10 mm precipitation threshold between 1200 UTC and 1800 UTC (Figure 8a). Concerning the categorical statistical measures, the ZTD assimilation results in the increase of the MAE by ~1 mm (~12%) for the highest rainfall threshold from 0000–0600 UTC (Figure 9a). Statistically significant reductions by 5.4% (90% confidence level) and 8.5% (95% confidence level) provided by the ZTD experiment are found for the precipitation intervals [2,5) and [10,20) from 0000–0600 UTC and 0600–1200 UTC, respectively (Figure 9). MAE is also decreased by 8.4% (statistically significant at the 95% confidence interval) during the ZTD simulation when rainfall is greater than 20 mm in the 6 h forecast from 1200–1800 UTC (Figure 10a). For most of the lower than 20 mm precipitation thresholds, the ZTD simulation between 1200 and 0000 UTC provides improvement. The MB values show that the WRF model overestimates the three lowest rainfall thresholds, with error magnitudes lower than ~4 mm for all 6 h intervals. In contrast, the precipitation amounts that are higher than 5 mm are underestimated by the model and the extent of under-prediction increases with the forecast lead time (Figures 9 and 10).

3.2. Analysis of Selected Case Studies
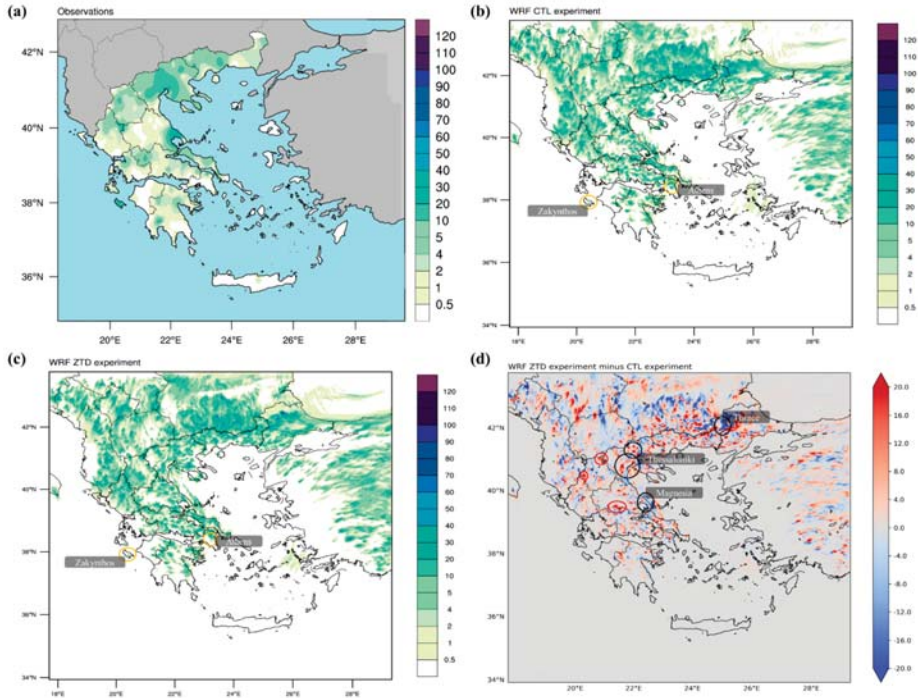
3.2.1. 10 May 2018

Figure 11a–b illustrate the atmospheric conditions at 0600 UTC, as simulated by the CTL experiment. As seen in the 500 hPa geopotential height and wind (Figure 11a), Greece is affected by a cyclonic atmospheric circulation in the middle troposphere, with a cut-off low over Western and central Greece. The cold air pool aloft is accompanied by a surface low over the central part of the country (Figure 11b). These synoptic conditions typically occur during late spring in Greece, producing strong atmospheric instability, which in turn results in intense convective activity. Figure A1 shows that, when ZTD data are assimilated into the WRF model, the modeled circulation pattern is the same as in the CTL experiment (Figure 11a–b). However, a discrete displacement of the 500 hPa geopotential height and sea level pressure gradients is evident, demonstrating the modification of the initial conditions by the ZTD assimilation at the time of the ZTD experiment initialization (0600 UTC). Since ZTD is related to precipitable water (PW), the initial conditions alteration at 0600 UTC is also evident for the modeled PW, as shown in Figure 11c. The initial PW differences between the conducted experiments emerge in locations where GNSS stations are situated (Figures 11c and 1b).
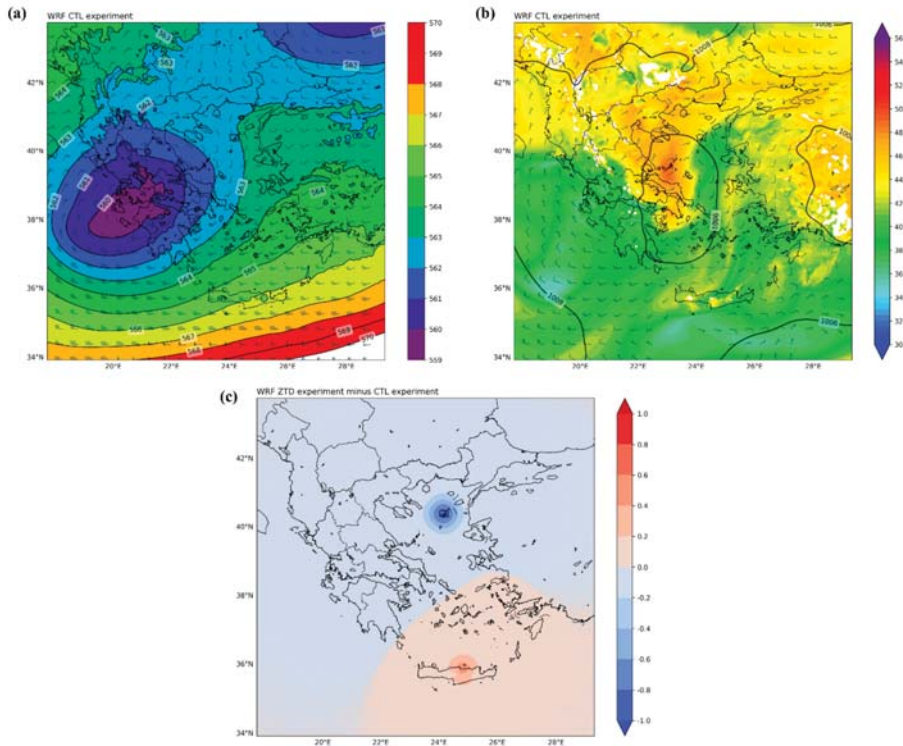
Figure 11. (**a**) WRF CTL simulated 500 hPa geopotential height (shading and contours) and wind barbs, (**b**) sea level pressure (contours) and 850 hPa equivalent potential temperature (shading) and wind barbs, and (**c**) PW differences between the CTL and ZTD experiments on 10 May 2018 at 0600 UTC.

The modification of the initial conditions of the ZTD experiment, due to the ZTD data assimilation, leads to differences in precipitation forecasts, as illustrated in Figure 12. In particular, the ZTD simulation (Figure 12c) improves the reproduction of the observed daily precipitation (Figure 12a) intensity and spatial distribution over the two high-rainfall regions, namely Thessaloniki and Magnesia. As presented in Figure 12d, with the black circles, higher precipitation amounts are simulated by the ZTD experiment over the city of Thessaloniki, as well as over the Western and Northern parts of the region. The ZTD assimilation also leads to greater (lower) daily rainfall values that are closer to the observations over the Northern (Southern) part of the Magnesia region (highlighted by the black circle in Figure 12d). Moreover, the assimilation of ZTD data into the WRF model results in the reduction of the significant precipitation overestimation produced by the CTL simulation over Northeast Thrace (Figure 12d). However, the ZTD experiment simulates considerably higher daily precipitation accumulations compared to the observations and CTL experiment over some regions, such as those highlighted by red circles in Figure 12d (e.g., central Greece). Overall, the model performs adequately in capturing the observed rainfall distribution over Greece, except the Zakynthos and Athens regions (orange circles in Figure 12b–c), where both CTL and ZTD simulations fail to reproduce the precipitation amounts, even though it is worth mentioning that the ZTD experiment reduces the geographical extent of the false forecast over the Athens area. The improvement in the reproduction of the observed rainfall when ZTD data are assimilated into the WRF model is also evident in the diurnal precipitation cycle analysis. As indicatively shown in Figure A2, most of the daily rain over the Thessaloniki region was observed between 0600 UTC and 1200 UTC, with the ZTD experiment

simulating notably higher precipitation amounts (Figure A2d), which is in better agreement with the observations (Figure A2a).

Figure 12. Daily (18 h) precipitation from (**a**) observations, (**b**) CTL and (**c**) ZTD simulation, and (**d**) differences between the two experiments for the 10 May 2018 event.

3.2.2. 18 November 2018

During the selected wet period event, Greece was affected by a deep surface low-pressure system. As shown in the sea-level pressure and 850 hPa equivalent potential temperature and wind fields (Figure 13b), the surface low is located over the Southwest Ionian Sea at 0000 UTC on 18 November, resulting in very strong Southerly winds. During the next hours, as the surface low is moving Northeast, the Southwesterly flow advects warm and moist air over Greece (Figure A3a–b). As in the case of the dry season event (see Section 3.2.1), both CTL and ZTD experiments simulate the same atmospheric conditions with slight differences in the 500 hPa geopotential height and sea level pressure gradients (Figure A4) at the ZTD experiment initialization time (0000 UTC). Again, the impact introduced by the ZTD data assimilation is clearly evident in the initial PW field at 0000 UTC (Figure 13c). The PW differences at 0000 UTC between the two experiments arise again over locations where GNSS stations are found (Figure 1b), but a more complicated pattern is evident in this case study.
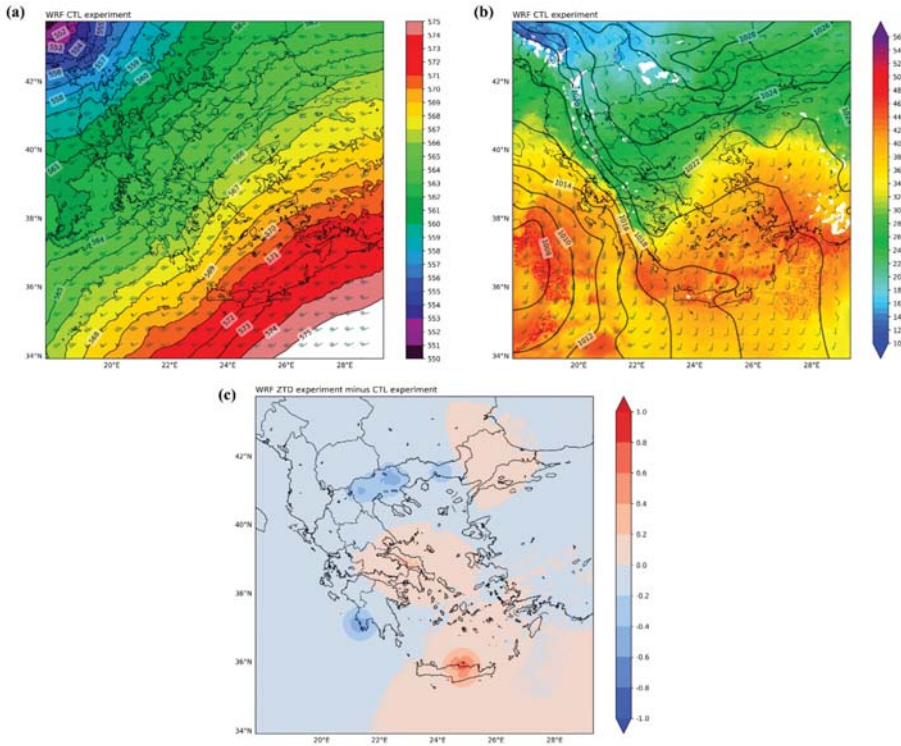
Figure 13. (**a**) WRF CTL simulated 500 hPa geopotential height (shading and contours) and wind barbs, (**b**) sea level pressure (contours) and 850 hPa equivalent potential temperature (shading) and wind barbs, and (**c**) PW differences between the CTL and ZTD experiments on 18 November 2018 at 0000 UTC.

Similarly, the 24 h modeled precipitation differences between the CTL and ZTD simulations arising from the modification of the initial conditions of the ZTD experiment due to the ZTD assimilation are more complex and widespread during this event (Figure 14d). The most noticeable differences are found over Western Greece, where the highest rainfall amounts were observed (Figure 14a). In particular, as shown in Figure 14d with black circles, the ZTD experiment produces lower (higher) precipitation over Kefalonia Island, the Aitolokarnania region, and central Peloponnese (the center of Zakynthos island, and Southwest and Southeast Peloponnese) compared to the CTL simulation, more accurately representing the observed rainfall. A more accurate reproduction of the precipitation observations by the ZTD experiment is also evident in Northwest Peloponnese, while in contrast, the ZTD assimilation leads to a more pronounced overestimation of the observed rainfall over central–North Peloponnese and Kythera Island, as indicated by the red circles in Figure 14d. Mixed results are found over the Epirus region, where the ZTD (CTL) experiment forecasts a rainfall pattern that is closer to the observed one in the central and Northern (Southern) part of the area. Overall, the model lacks the ability to capture the magnitude of observations, especially for intense precipitation, as shown by the notable underestimation (overestimation) of the observed rainfall over the area denoted by the orange circle in Chalkidiki (Magnesia) region in Figure 14b–c. However, the spatial distribution of the observed precipitation is well captured by both CTL and ZTD simulations. The above findings are also evident when examining the diurnal precipitation cycle in 6 h intervals and highlight the good performance of the WRF model in reproducing the temporal distribution of the observed rainfall, especially during

the ZTD experiment. Indicatively, both simulations (Figure A5b–c) adequately replicate the observed spatial pattern of the precipitation between 0600 and 1200 UTC (Figure A5a), with the ZTD experiment improving the rainfall amounts forecasts over North Epirus, Aitolokarnania and Peloponnese regions, and Kefalonia and Zakynthos Islands (Figure A5d).
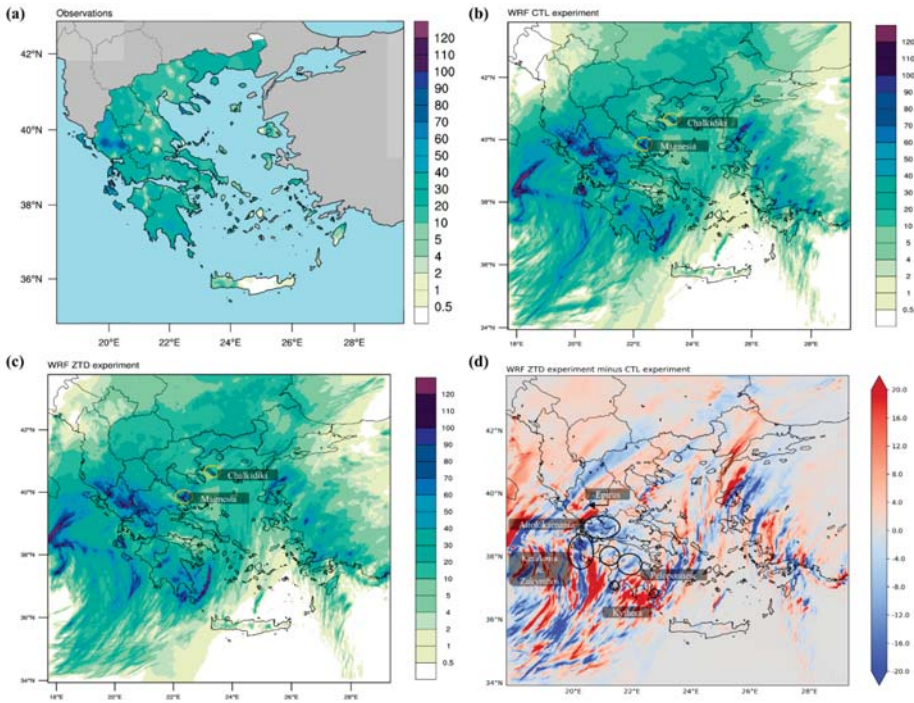
Figure 14. Daily (24 h) precipitation from (**a**) observations, (**b**) CTL and (**c**) ZTD simulation, and (**d**) differences between the two experiments for the 18 November 2018 event.

4. Discussion

The results presented above demonstrate a beneficial impact of ZTD assimilation on the daily precipitation forecasts over Greece during both dry and wet periods (Figures 3 and 4). This impact is more profound for heavy precipitation (>20 mm), for which statistically significant (at least at the 90% confidence level) improvements are provided by the ZTD assimilation concerning the probability of detection, the false alarm ratio, the quality of forecasts (ETS), and the magnitude of errors (MAE). This finding is of great importance as higher rainfall amounts are associated with more severe impacts, especially over the Mediterranean regions [20]. In general, the positive impacts of ZTD assimilation are more evident during the dry period. However, notable and statistically significant enhancements in the WRF model's performance are also found during the ZTD experiment for both 24 h (Figures 3 and 4) and 6 h precipitation (Figures 7–10). This fact reflects a significant added value of ZTD assimilation during wet season rainfall events, in contrast with previous studies, which concluded that when the atmospheric state is well described during large synoptic-scale cases, no further improvement of the forecast accuracy can be achieved by assimilating ZTD observations (e.g., Boniface et al. [24]). Moreover, the ZTD assimilation substantially advances the overall model performance in terms of qualitative and quantitative rainfall forecasting between 0600 and 1800 UTC in the dry season (Figures 5 and 6). Especially from 1200–1800 UTC, statistically significant (at the 95% confidence interval) higher POD, ETS, and MAE values, compared to the CTL experiment, are found across all precipitation thresholds.

This is a key outcome considering that convective precipitation in the dry season mainly occurs during the examined time window and its forecasting is challenging, as NWP models lack the ability to resolve small-scale convective circulation [68,69]. In contrast, NWP models are more capable in dealing with large-scale dynamically-driven precipitation systems during the wet season [70]. This is evident in the current study because the WRF model, overall, performs better in the wet compared to the dry period. The case study analysis shows that ZTD assimilation provides a more accurate representation of the observed precipitation geographical extent in terms of 24 h and 6 h rainfall (Figures 12 and A2 in the dry period and Figures 14 and A5 in the wet period). It also highlights that the overall WRF performance improvement in rainfall forecasting under the ZTD experiment is due to the modification of the model's initial conditions at the time of the experiment's initialization through the assimilation of ZTD observations. The introduction of ZTD data assimilation especially affects the initial PW field, leading to noticeable differences between the conducted experiments, which emerge in locations where GNSS stations are found. This is evident in Figures 11c and 13c concerning the selected case studies, as well as in Figures A6–A13, illustrating the PW differences between the CTL and ZTD simulations at the ZTD experiment initialization time for each day of the rest of the rainfall events examined in the present study.

5. Conclusions

The present work focuses on investigating the impact of assimilating GNSS ZTD observations on the WRF model precipitation forecast skill. The conducted work is the first attempt at applying ZTD data assimilation into a regional atmospheric model over Greece. The evaluation of the impact of ZTD assimilation is performed for seven high-impact rainfall episodes that occurred during the dry (four events) and wet period (three events) of 2018. A substantial effort was devoted to the pre-processing of ZTD observations in order to qualify their adequacy for data assimilation. Then, two sets of model experiments (CTL and ZTD) were performed and the model predicted rainfall was verified against observations, which were collected from over 330 rain gauges provided by the weather monitoring network of the National Observatory of Athens.

The qualitative and quantitative statistical evaluation provides substantial findings on how the assimilation of ZTD observations into the WRF model affects its performance:

- The ZTD assimilation results into statistically significant (a)more accurate reproduction of the occurrence of the observed precipitation (higher POD by 10.6% and 3.9% in the dry and wet season, respectively), (b) reduction of the false forecasts (lower FAR by 7.7% and 16.9% in the dry and wet season, respectively), (c) better prediction quality (greater ETS by 11.4% and 8.1% in the dry and wet season, respectively), and (d) decrease in the magnitude of errors (lower MAE by 6.3% and 5.5% in the dry and wet season, respectively), compared to the CTL configuration, for intense rainfall (>20 mm). This outcome is of great importance when considering that heavy precipitation amounts are associated with greater impacts and may be poorly forecasted.
- The overall model performance enhancement in rainfall forecasting during the ZTD experiment is more evident in the dry season. However, the assimilation of ZTDs also leads to notable and statistically significant improvements during the wet period, indicating, in contrast to past studies, that it can provide a positive influence under large-scale synoptic conditions.
- The introduction of ZTDs into the WRF model induces statistically significant improvements in precipitation forecasts, especially for above 20 mm 6 h accumulations, during the time window (0600 to 1800 UTC) of convective rain occurrence in the dry season. This finding is essential as correctly forecasting high precipitation convective events during the dry period is crucial in issuing improved severe rainfall warnings.

The detailed presentation of two case studies (on 10 May and 18 November, 2018) show that the WRF model is capable of satisfactorily capturing the spatial and temporal distribution of the observed rainfall, with the ZTD simulation providing results that are closer to the observations. It also reveals that the

precipitation forecast improvement under the ZTD experiment is induced by the initial conditions' modification at the experiment's initialization time, especially affecting the initial PW field, due to the ZTD assimilation.

The above findings highlight the beneficial impact of assimilating ZTD observations into high-resolution regional scale weather forecasting systems. It is the authors' intention to investigate the impact of variational ZTD bias correction techniques and of ZTD spatiotemporal error correlations in the assimilation framework.

Author Contributions: Conceptualization, C.G., V.K., K.L., T.M.G., C.P.; Data curation, C.G., C.P.; Formal analysis, C.G.; Funding acquisition, V.K., K.L., C.P.; Investigation, C.G.; Methodology, C.G., V.K., K.L., T.M.G.; Resources, V.K., K.L., C.P.; Software, C.G., T.M.G.; Supervision, V.K., K.L, T.M.G; Validation, C.G., V.K., K.L., T.M.G; Visualization, C.G.; Writing – original draft, C.G.; Writing – review & editing, C.G., V.K., K.L, T.M.G., C.P. All authors have read and agreed to the published version of the manuscript.

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

Appendix A

Figure A1. As in Figure 11a,b, but for the ZTD experiment.
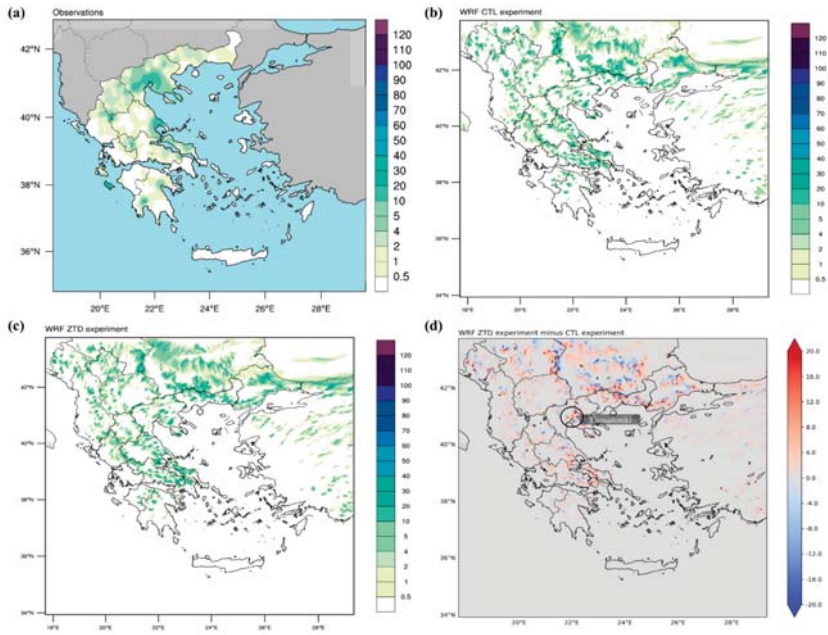
Figure A2. As in Figure 12a–d, but for the 6-h accumulated precipitation at 10 May 2018 1200 UTC.
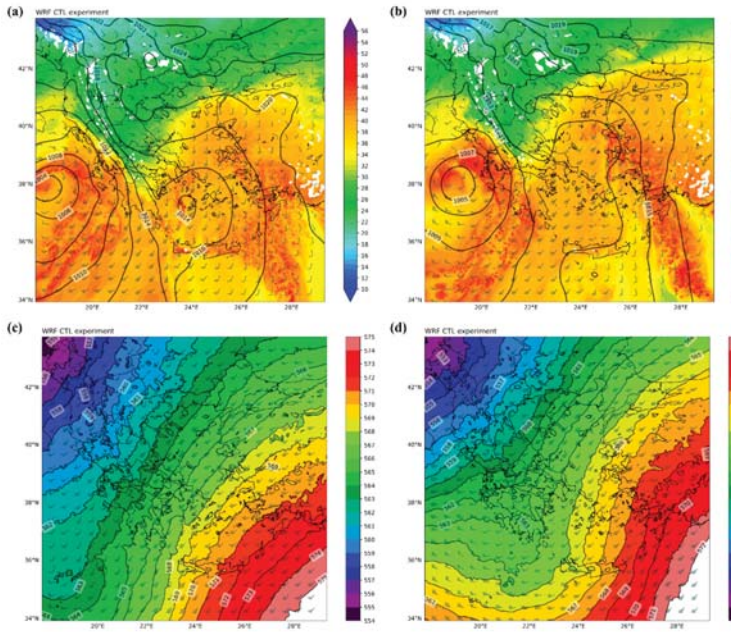
Figure A3. As in Figure 13a,b, but for 18 November 2018 (**a**,**c**) 0900UC and (**b**,**d**) 1500 UTC.
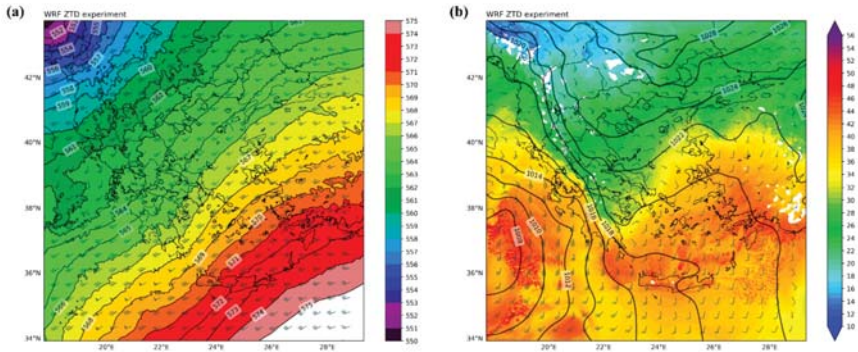
Figure A4. As in Figure 13a,b, but for the ZTD experiment.
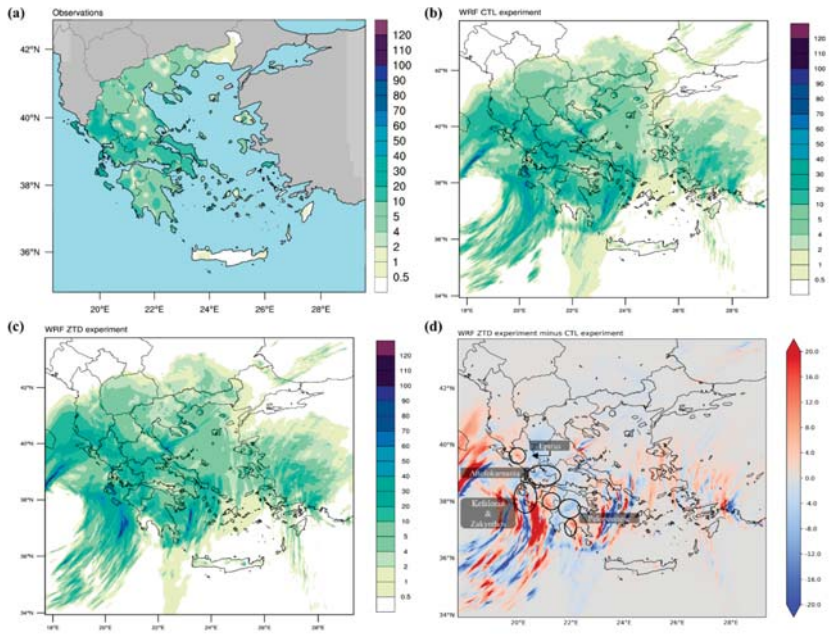
Figure A5. As in Figure 14a–d, but for the 6-h accumulated precipitation at 18 November 2018 1200 UTC.
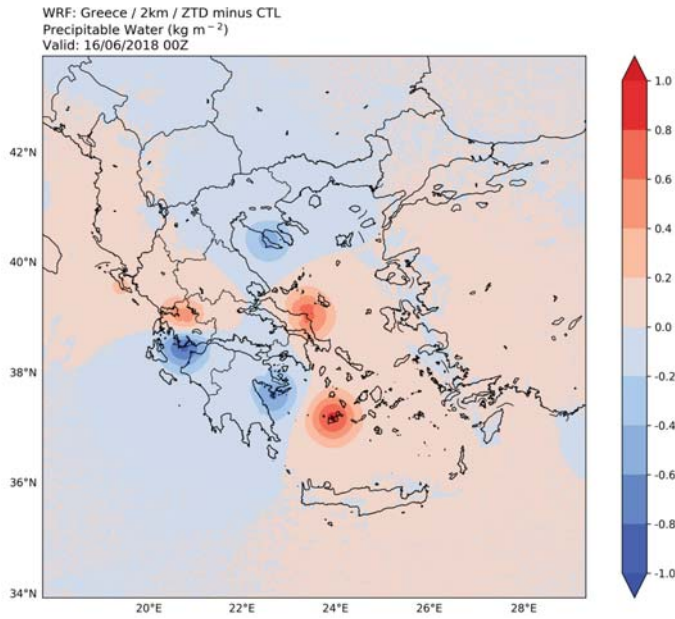
Figure A6. As in Figures 11c and 13c, but for 16 June 2018 0000 UTC.
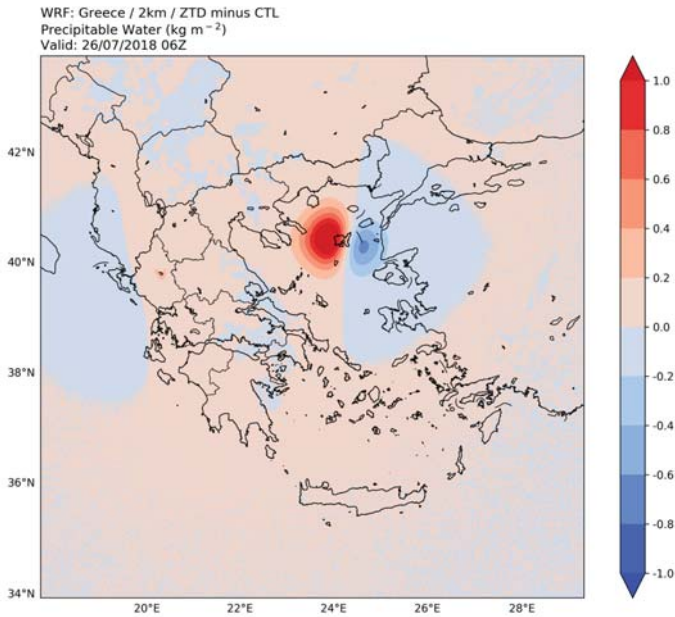
Figure A7. As in Figures 11c and 13c, but for 26 July 2018 0600 UTC.
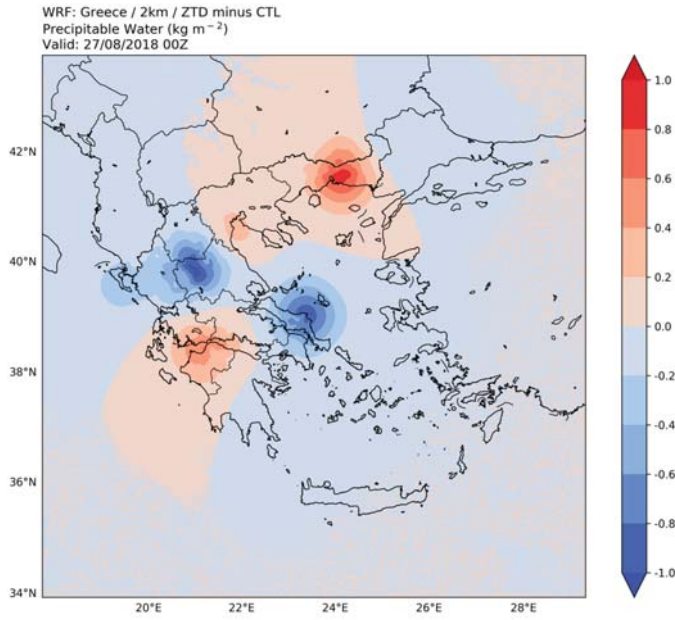
Figure A8. As in Figures 11c and 13c, but for 27 August 2018 0000 UTC.
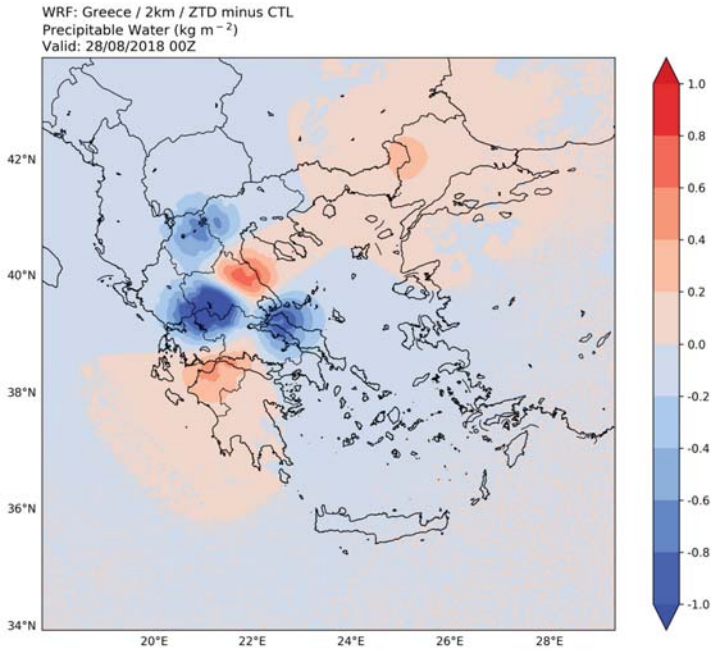
Figure A9. As in Figures 11c and 13c, but for 28 August 2018 0000 UTC.
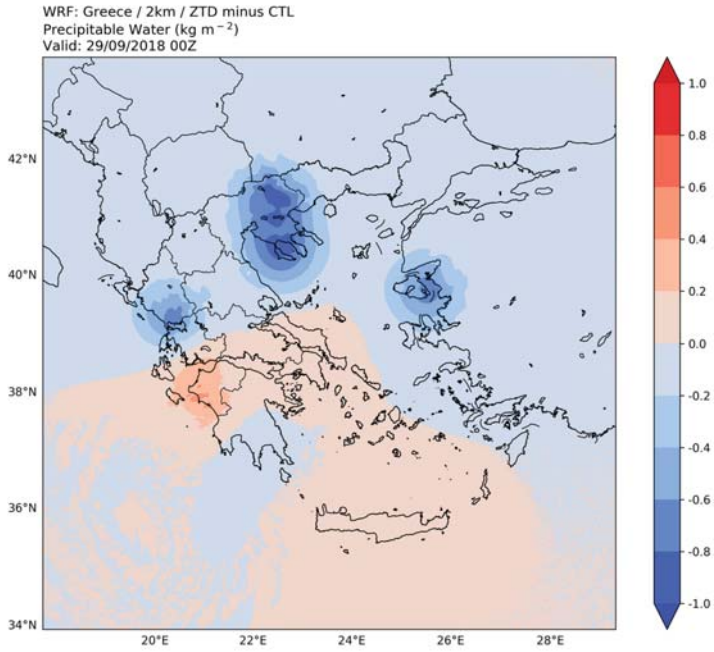
Figure A10. As in Figures 11c and 13c, but for 29 September 2018 0000 UTC.
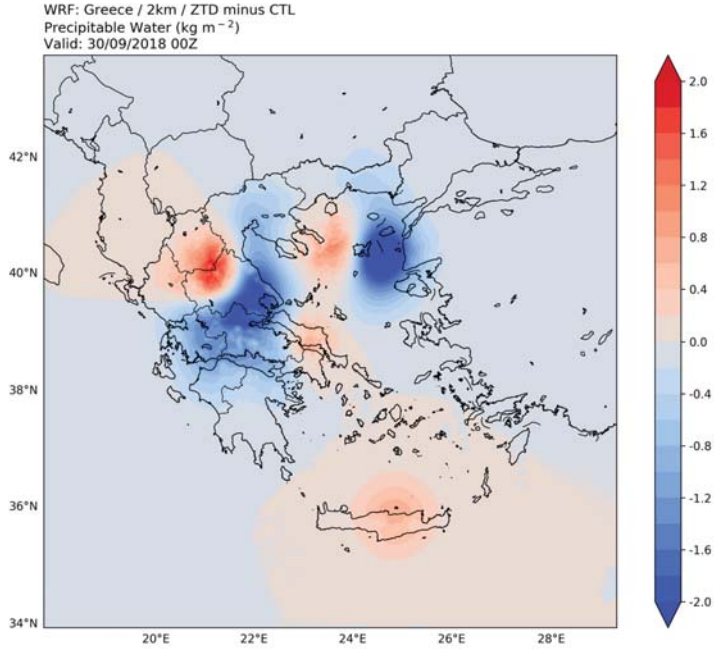
Figure A11. As in Figures 11c and 13c, but for 30 September 2018 0000 UTC.
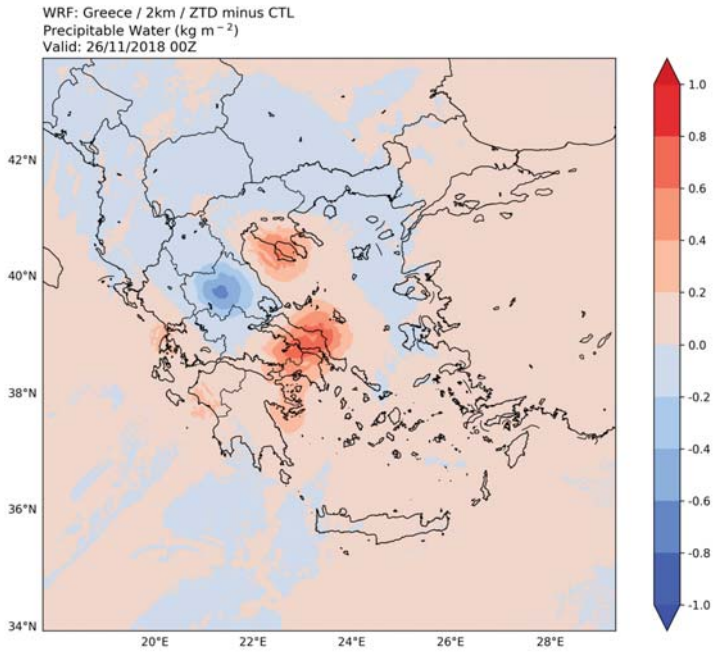
Figure A12. As in Figures 11c and 13c, but for 26 November 2018 0000 UTC.
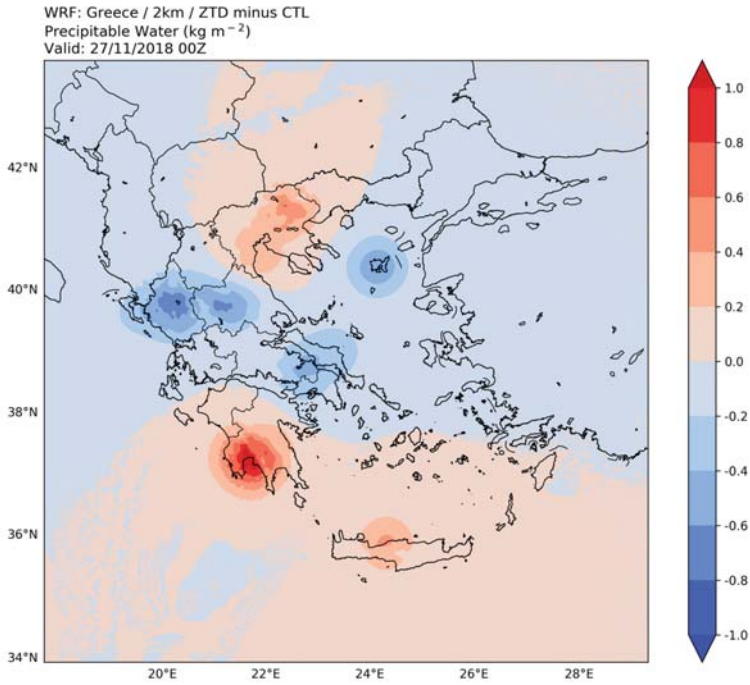
Figure A13. As in Figures 11c and 13c, but for 27 November 2018 0000 UTC.

References

1. Sabatini, R.; Moore, T.; Ramasamy, S. Global navigation satellite systems performance analysis and augmentation strategies in aviation. *Prog. Aerosp. Sci.* **2017**, *95*, 45–98. [CrossRef]
2. Kubo, N.; Higuchi, M.; Takasu, T.; Yamamoto, H. Performance evaluation of GNSS-based railway applications. In Proceedings of the 2015 International Association of Institutes of Navigation World Congress (IAIN), Prague, Czech Republic, 20–23 October 2015; pp. 1–8.
3. Molina, P.; Colomina, I.; Vitoria, T.; Silva, P.F.; Skaloud, J.; Kornus, W.; Prades, R.; Aguilera, C. Searching Lost People with Uavs: The System and Results of the Close-Search Project. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *39*, 441–446. [CrossRef]
4. Kahveci, M. Contribution of GNSS in precision agriculture. In Proceedings of the 2017 8th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, Turkey, 19–22 June 2017; pp. 513–516.
5. Ostolaza, J.; Lera, J.J.; Pérez, D.; Cueto-Felgueroso, G.; Cueto, M.; Cezón, A.; Fernández, M.A.; López, M.; Hill, D.; Boissinot, V.; et al. Maritime Trials in Europe and Africa Using GNSS-based Enhanced Systems. In Proceedings of the 29th International Technical Meeting of the Satellite Division of The Institute of Navigation, Portland, OR, USA, 12–16 September 2016.
6. Guerova, G.; Jones, J.; Douša, J.; Dick, G.; De Haan, S.; Pottiaux, E.; Bock, O.; Pacione, R.; Elgered, G.; Vedel, H.; et al. Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. *Atmos. Meas. Tech.* **2016**, *9*, 5385–5406. [CrossRef]
7. Bevis, M.; Businger, S.; Herring, T.A.; Rocken, C.; Anthes, R.A.; Ware, R.H. GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. Geophys. Res.* **2012**, *97*, 15787. [CrossRef]
8. Suparta, W.; Adnan, J.; Mohd Ali, M.A. Nowcasting the lightning activity in Peninsular Malaysia using the GPS PWV during the 2009 inter-monsoons. *Ann. Geophys.* **2014**, *57*, 0217.
9. Dousa, J.; Vaclavovic, P. Real-time zenith tropospheric delays in support of numerical weather prediction applications. *Adv. Space Res.* **2014**, *53*, 1347–1358. [CrossRef]
10. Hanna, N.; Trzcina, E.; Möller, G.; Rohm, W.; Weber, R. Assimilation of GNSS tomography products into WRF using radio occultation data assimilation operator. *Atmos. Meas. Tech. Discuss.* **2019**, *12*, 4829–4848. [CrossRef]
11. Simeonov, T.; Sidorov, D.; Teferle, F.N.; Milev, G.; Guerova, G. Evaluation of IWV from the numerical weather prediction WRF model with PPP GNSS processing for Bulgaria. *Atmos. Meas. Tech. Discuss.* **2016**. [CrossRef]
12. Manning, T.; Zhang, K.; Rohm, W.; Choy, S. Detecting Severe Weather using GPS Tomography: An Australian Case Study. *J. Glob. Position. Syst.* **2012**, *11*, 58–70. [CrossRef]
13. Pikridas, C. Monitoring Climate Changes on Small Scale Networks Using Ground Based GPS and Meteorological Data. *Artif. Satell.* **2014**, *49*, 125–135. [CrossRef]
14. Pikridas, C. The use of GNSS tropospheric products for climate monitoring. A case study in the area of Ioannina Northwestern Greece. *South East. Eur. J. Earth Obs. Geomat.* **2015**, *4*, 81–90.
15. Pikridas, C.; Fotiou, A.; Bitharis, S.; Karolos, I.; Balidakis, K. BeRTISS: Balkan-Mediterranean real time severe weather service. The case of Greece. In Proceedings of the 5th International Conference on Civil Protection & New Technologies, Kozani, Greece, 31 October–3 November 2018.
16. Arriola, J.S.; Lindskog, M.; Thorsteinsson, S.; Bojarova, J. Variational bias correction of GNSS ZTD in the HARMONIE modeling system. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 1259–1276. [CrossRef]
17. E-GVAP. Available online: http://egvap.dmi.dk/ (accessed on 24 January 2019).
18. Poli, P.; Moll, P.; Rabier, F.; Desrozier, G.; Chapnik, B.; Berre, L.; Healy, S.B.; Andersson, E.; El Guelai, F.Z. Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. *J. Geophys. Res. Atmos.* **2007**, *112*, 47. [CrossRef]
19. Yan, X.; Ducrocq, V.; Jaubert, G.; Brousseau, P.; Poli, P.; Champollion, C.; Flamant, C.; Boniface, K. The benefit of GPS zenith delay assimilation to high-resolution quantitative precipitation forecasts: A case-study from COPS IOP 9. *Q. J. R. Meteorol. Soc.* **2009**, *135*, 1788–1800. [CrossRef]
20. Yan, X.; Ducrocq, V.; Poli, P.; Hakam, M.; Jaubert, G.; Walpersdorf, A. Impact of GPS zenith delay assimilation on convective-scale prediction of Mediterranean heavy rainfall. *J. Geophys. Res. Atmos.* **2009**, *114*, 20. [CrossRef]

21. Eresmaa, R.; Salonen, K.; Järvinen, H. An observing-system experiment with ground-based GPS zenith total delay data using HIRLAM 3D-Var in the absence of satellite data. *Q. J. R. Meteorol. Soc.* **2010**, *136*, 1289–1300. [CrossRef]

22. Bennitt, G.V.; Jupp, A. Operational Assimilation of GPS Zenith Total Delay Observations into the Met Office Numerical Weather Prediction Models. *Mon. Weather Rev.* **2012**, *140*, 2706–2719. [CrossRef]

23. Schwitalla, T.; Bauer, H.-S.; Wulfmeyer, V.; Aoshima, F. High-resolution simulation over central Europe: Assimilation experiments during COPS IOP 9c. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 156–175. [CrossRef]

24. Boniface, K.; Ducrocq, V.; Jaubert, G.; Yan, X.; Brousseau, P.; Masson, F.; Champollion, C.; Chéry, J.; Doerflinger, E. Impact of high-resolution data assimilation of GPS zenith delay on Mediterranean heavy rainfall forecasting. *Ann. Geophys.* **2009**, *27*, 2739–2753. [CrossRef]

25. Bennitt, G.V.; Johnson, H.R.; Weston, P.P.; Jones, J.; Pottiaux, E. An assessment of ground-based GNSS Zenith Total Delay observation errors and their correlations using the Met Office UKV model. *Q. J. R. Meteorol. Soc.* **2017**, *143*, 2436–2447. [CrossRef]

26. Macpherson, S.; Laroche, S. Estimation of ground-based GNSS Zenith Total Delay temporal observation error correlations using data from the NOAA and E-GVAP networks. *Q. J. R. Meteorol. Soc.* **2019**, *145*, 513–529. [CrossRef]

27. Mile, M.; Benáček, P.; Rózsa, S. The use of GNSS zenith total delays in operational AROME/Hungary 3D-Var over a central European domain. *Atmos. Meas. Tech.* **2019**, *12*, 1569–1579. [CrossRef]

28. Rohm, W.; Guzikowski, J.; Wilgan, K.; Kryza, M. 4DVAR assimilation of GNSS zenith path delays and precipitable water into a numerical weather prediction model WRF. *Atmos. Meas. Tech.* **2019**, *12*, 345–361. [CrossRef]

29. Lagouvardos, K.; Kotroni, V.; Bezes, A.; Koletsis, I.; Kopania, T.; Lykoudis, S.; Mazarakis, N.; Papagiannaki, K.; Vougioukas, S. The automatic weather stations NOANN network of the National Observatory of Athens: Operation and database. *Geosci. Data J.* **2017**, *4*, 4–16. [CrossRef]

30. Papagiannaki, K.; Lagouvardos, K.; Kotroni, V. A database of high-impact weather events in Greece: A descriptive impact analysis for the period 2001–2011. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 727–736. [CrossRef]

31. Michaelides, S.; Karacostas, T.; Sánchez, J.L.; Retalis, A.; Pytharoulis, I.; Homar, V.; Romero, R.; Zanis, P.; Giannakopoulos, C.; Bühl, J.; et al. Reviews and perspectives of high impact atmospheric processes in the Mediterranean. *Atmos. Res.* **2018**, *208*, 4–44. [CrossRef]

32. Galanaki, E.; Kotroni, V.; Lagouvardos, K.; Argiriou, A. A ten-year analysis of cloud-to-ground lightning activity over the Eastern Mediterranean region. *Atmos. Res.* **2015**, *166*, 213–222. [CrossRef]

33. Kotroni, V.; Lagouvardos, K. Lightning in the Mediterranean and its relation with sea-surface temperature. *Environ. Res. Lett.* **2016**, *11*, 34006. [CrossRef]

34. Lagouvardos, K.; Kotroni, V.; Dobricic, S.; Nickovic, S.; Kallos, G. The storm of October 21–22, 1994, over Greece: Observations and model results. *J. Geophys. Res. Atmos.* **1996**, *101*, 26217–26226. [CrossRef]

35. Lagouvardos, K.; Kotroni, V.; Defer, E. The 21–22 January 2004 explosive cyclogenesis over the Aegean Sea: Observations and model analysis. *Q. J. R. Meteorol. Soc.* **2007**, *133*, 1519–1531. [CrossRef]

36. Kotroni, V.; Lagouvardos, K.; Kallos, G.; Ziakopoulos, D. Severe flooding over central and southern greece associated with pre-cold frontal orographic lifting. *Q. J. R. Meteorol. Soc.* **1999**, *125*, 967–991. [CrossRef]

37. Lagouvardos, K.; Kotroni, V.; Nickovic, S.; Jovic, D.; Kallos, G.; Tremback, C.J. Observations and model simulations of a winter sub-synoptic vortex over the central Mediterranean. *Meteorol. Appl.* **1999**, *6*, 371–383. [CrossRef]

38. Fita, L.; Flaounas, E. Medicanes as subtropical cyclones: The December 2005 case from the perspective of surface pressure tendency diagnostics and atmospheric water budget. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 1028–1044. [CrossRef]

39. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Liu, Z.; Berner, J.; Wang, W.; Powers, J.G.; Duda, M.G.; Barker, D.; et al. *A Description of the Advanced Research WRF Model Version 4*; NCAR: Boulder, CO, USA, 2008.

40. Galanaki, E.; Lagouvardos, K.; Kotroni, V.; Giannaros, T.M.; Giannaros, C. Calibration and evaluation of WRF-Hydro performance at two drainage basins in the region of Attica, Greece. In Proceedings of the EMS Annual Meeting, Copenhagen, Denmark, 9–13 September 2019.

41. Giannaros, T.M.; Kotroni, V.; Lagouvardos, K. IRIS—Rapid response fire spread forecasting system: Development, calibration and evaluation. *Agric. For. Meteorol.* **2019**, *279*, 107745. [CrossRef]

42. Thompson, G.; Field, P.R.; Rasmussen, R.M.; Hall, W.D. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.* **2008**, *136*, 5095–5115. [CrossRef]

43. Kain, J.S. The Kain–Fritsch Convective Parameterization: An Update. *J. Appl. Meteorol.* **2004**, *43*, 170–181. [CrossRef]

44. Hong, S.; Lim, J. The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteorol. Soc.* **2006**, *42*, 129–151.

45. Jiménez, P.A.; Dudhia, J.; González-Rouco, J.F.; Navarro, J.; Montávez, J.P.; García-Bustamante, E. A Revised Scheme for the WRF Surface Layer Formulation. *Mon. Weather Rev.* **2012**, *140*, 898–918. [CrossRef]

46. Tewari, M.; Chen, F.; Wang, W.; Dudhia, J.; LeMone, M.A.; Mitchell, K.; Ek, M.; Gayno, G.; Wegiel, J.; Cuenca, R. Implementation and verification of the united NOAH land surface model in the WRF model. In Proceedings of the 20th Conference of Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction, American Meteorological Society, Seattle, WA, USA, 14 January 2004.

47. Iacono, M.J.; Delamere, J.S.; Mlawer, E.J.; Shephard, M.W.; Clough, S.A.; Collins, W.D. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res. Atmos.* **2008**, *113*, 2–9. [CrossRef]

48. Barker, D.M.; Huang, W.; Guo, Y.-R.; Bourgeois, A.J.; Xiao, Q.N. A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results. *Mon. Weather Rev.* **2004**, *132*, 897–914. [CrossRef]

49. Huang, X.-Y.; Xiao, Q.; Barker, D.M.; Zhang, X.; Michalakes, J.; Huang, W.; Henderson, T.; Bray, J.; Chen, Y.; Ma, Z.; et al. Four-Dimensional Variational Data Assimilation for WRF: Formulation and Preliminary Results. *Mon. Weather Rev.* **2009**, *137*, 299–314. [CrossRef]

50. Mazzarella, V.; Maiello, I.; Capozzi, V.; Budillon, G.; Ferretti, R. Comparison between 3D-Var and 4D-Var data assimilation methods for the simulation of a heavy rainfall case in central Italy. *Adv. Sci. Res.* **2017**, *14*, 271–278. [CrossRef]

51. Waller, J.A.; Dance, S.L.; Lawless, A.S.; Nichols, N.K. Estimating correlated observation error statistics using an ensemble transform Kalman filter. *Tellus A Dyn. Meteorol. Oceanogr.* **2014**, *66*, 23294. [CrossRef]

52. Parrish, D.F.; Derber, J.C. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Weather Rev.* **1992**, *120*, 1747–1763. [CrossRef]

53. Yucel, I.; Onen, A. Evaluating a mesoscale atmosphere model and a satellite-based algorithm in estimating extreme rainfall events in northwestern Turkey. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 611–624. [CrossRef]

54. EUREF. Available online: http://www.euref.eu/ (accessed on 24 January 2019).

55. Pikridas, C.; Katsougiannopoulos, S.; Zinas, N. A comparative study of zenith tropospheric delay and precipitable water vapor estimates using scientific GPS processing software and web based automated PPP service. *Acta Geod. Geophys.* **2014**, *49*, 177–188. [CrossRef]

56. Fotiou, A.; Pikridas, C.; Rossikopoulos, D.; Katsougiannopoulos, S.; Bitharis, S.; Karolos, I. Geodetic activities of GNSS QC research team of AUTh. In Proceedings of the EUREF 2016 Symposium, San Sebastian, Spain, 25–27 May 2016.

57. Katsougiannopoulos, S.; Pikridas, C.; Zinas, C.; Karolos, I.; Bitharis, S. Near Real Time graphical representation of Tropospheric and Positioning products. In Proceedings of the COST-ES1206 Final Workshop, Noordwijk, The Netherlands, 21–23 February 2017.

58. Dach, R.; Lutz, S.; Walser, P.; Fridez, P. *Bernese GNSS Sofware Version 5.2*; University of Bern, Bern Open Publishing: Bern, Germany, 2015.

59. Pikridas, C.; Fotiou, A.; Karolos, I.; Bitharis, S. First Results of BeRTISS Project. In Proceedings of the EGU General Assembly, Vienna, Austria, 8–13 April 2018.

60. Bosy, J.; Kaplon, J.; Rohm, W.; Sierny, J.; Hadas, T. Near real-time estimation of water vapour in the troposphere using ground GNSS and the meteorological data. *Ann. Geophys.* **2012**, *30*, 1379–1391. [CrossRef]

61. Dymarska, N.; Rohm, W.; Sierny, J.; Kapłon, J.; Kubik, T.; Kryza, M.; Jutarski, J.; Gierczak, J.; Kosierb, R. An assessment of the quality of near-real time GNSS observations as a potential data source for meteorology. *Meteorol. Hydrol. Water Manag.* **2017**, *5*, 3–13. [CrossRef]

62. Wilgan, K.; Rohm, W.; Bosy, J. Multi-observation meteorological and GNSS data comparison with Numerical Weather Prediction model. *Atmos. Res.* **2015**, *156*, 29–42. [CrossRef]

63. Offiler, D. *EIG EUMETNET GNSS Water Vapour Programme (E-GVAP-II)*; Product Requirements Document; Met Office: Yassett, UK, 2010.

64. Giannaros, T.M.; Kotroni, V.; Lagouvardos, K. WRF-LTNGDA: A lightning data assimilation technique implemented in the WRF model for improving precipitation forecasts. *Environ. Model. Softw.* **2016**, *76*, 54–68. [CrossRef]

65. Dafis, S.; Lagouvardos, K.; Kotroni, V.; Giannaros, T.M.; Bartzokas, A. Observational and modeling study of a mesoscale convective system during the HyMeX—SOP1. *Atmos. Res.* **2017**, *187*, 1–15. [CrossRef]

66. Dafis, S.; Fierro, A.; Giannaros, T.M.; Kotroni, V.; Lagouvardos, K.; Mansell, E. Performance Evaluation of an Explicit Lightning Forecasting System. *J. Geophys. Res. Atmos.* **2018**, *123*, 5130–5148. [CrossRef]

67. Sikder, M.S.; Hossain, F. Sensitivity of initial-condition and cloud microphysics to the forecasting of monsoon rainfall in South Asia. *Meteorol. Appl.* **2018**, *25*, 493–509. [CrossRef]

68. Mazarakis, N.; Kotroni, V.; Lagouvardos, K.; Argiriou, A.A. The sensitivity of numerical forecasts to convective parameterization during the warm period and the use of lightning data as an indicator for convective occurrence. *Atmos. Res.* **2009**, *94*, 704–714. [CrossRef]

69. Sun, J.; Trier, S.B.; Xiao, Q.; Weisman, M.L.; Wang, H.; Ying, Z.; Xu, M.; Zhang, Y. Sensitivity of 0-12-h warm-season precipitation forecasts over the central United States to model initialization. *Weather Forecast.* **2012**, *27*, 832–855. [CrossRef]

70. Mercer, A.; Dyer, J.; Zhang, S. Warm-season thermodynamically-driven rainfall prediction with support vector machines. *Procedia Comput. Sci.* **2013**, *20*, 128–133. [CrossRef]

Article

Evaluation and Application of Satellite Precipitation Products in Studying the Summer Precipitation Variations over Taiwan

Wan-Ru Huang [1,*], Pin-Yi Liu [1], Ya-Hui Chang [1] and Chian-Yi Liu [2]

1   Department of Earth Sciences, National Taiwan Normal University, Taipei 11677, Taiwan;
     pinyiliu@ntnu.edu.tw (P.-Y.L.); yahuichang@ntnu.edu.tw (Y.-H.C.)
2   Center for Space and Remote Sensing Research, National Central University, Taoyuan 32001, Taiwan;
     cyliu@g.ncu.edu.tw
*   Correspondence: wrhuang@ntnu.edu.tw

Abstract: In March 2019, Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG)-Final v6 (hereafter IMERG6) was released, with data concerning precipitation dating back to June 2000. The National Aeronautics and Space Administration (NASA) has suggested that researchers use IMERG6 to replace the frequently used Tropical Rainfall Measuring Mission (TRMM)-3B42 v7 (hereafter TRMM7), which is expected to cease operation in December 2019. This study aims to evaluate the performance of IMERG6 and TRMM7 in depicting the variations of summer (June, July, and August) precipitation over Taiwan during the period 2000–2017. Data used for the comparison also includes IMERG-Final v5 (hereafter IMERG5) and Global Satellite Mapping of Precipitation for Global Precipitation Measurement (GSMaP)-Gauge v7 (hereafter GSMaP7) during the summers of 2014–2017. Capabilities to apply the four satellite precipitation products (SPPs) in studying summer connective afternoon rainfall (CAR) events, which are the most frequently observed weather patterns in Taiwan, are also examined. Our analyses show that when using more than 400 local rain-gauge observations as a reference base for comparison, IMERG6 outperforms TRMM7 quantitatively and qualitatively, more accurately depicting the variations of the summer precipitation over Taiwan at multiple timescales (including mean status, daily, interannual, and diurnal). IMERG6 also performs better than TRMM7 in capturing the characteristics of CAR activities in Taiwan. These findings highlight that using IMERG6 to replace TRMM7 adds value in studying the spatial-temporal variations of summer precipitation over Taiwan. Furthermore, the analyses also indicated that IMERG6 outperforms IMERG5 and GSMaP7 in the examination of most of the features of summer precipitation over Taiwan during 2014–2017.

Keywords: assessment; satellite precipitation; Taiwan

---

1. Introduction

The Tropical Rainfall Measuring Mission (TRMM) was launched in November 1997 [1], and its most frequently used product, TRMM-3B42 v7 (hereafter TRMM7), is expected to cease in December 2019 (https://pmm.nasa.gov/data-access/downloads/trmm). As an extension and enhancements on the TRMM data, the Global Precipitation Measurement (GPM) Core Satellite was launched in February 2014 [2]. After that, the satellite precipitation products (hereafter SPPs) of the GPM mission, including Integrated Multi-satellitE Retrievals for GPM (IMERG) [3] and Global Satellite Mapping of Precipitation for GPM (GSMaP) [4], were provided by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), respectively. In March 2019, NASA released the IMERG-Final v6 (hereafter IMERG6), which includes TRMM-era

data dating back to June 2000 and suggested that researchers use this for most research purposes (https://pmm.nasa.gov/data-access/downloads/trmm). However, the performance of these SPPs may depend heavily on location and season [5–12]. Thus, it is important to clarify which SPP (e.g., IMERG or GSMaP) is the most suitable product to replace TRMM7 for studies of precipitation changes over various regions and during various seasons [13–15]. A better understanding of the performance of SPPs [13–22] can benefit other studies where SPPs are required to examine issues that are related to precipitation (e.g., moisture budget, speed of the hydrological cycle, etc.).

Located in Asia, Taiwan (119.9°E–122.1°E, 21.8°N–25.5°N) is an island that is known for its complex terrain (Figure 1b). In view of earlier literature, few studies have evaluated the performance of IMERG6 [15] or other SPPs over Taiwan [15,23–26]. Recently, Huang et al. [24] showed that IMERG-Final v5 (hereafter IMERG5), which is the earlier version of IMERG6, can qualitatively illustrate the multiple timescale variations in precipitation over Taiwan in a similar manner to the local rain-gauge observations made during the period March 2014–February 2017, but the amount of the estimation is lower than that seen in the gauge observations. However, Huang et al. [24] did not compare IMERG5 with the other SPPs used to investigate the precipitation around Taiwan. In addition, it should be noted that a major change was made to the morphing scheme used in IMERG5 and IMERG6 [3]. In versions of IMERG up to and including v5, the vectors used to describe cloud motion were computed from geosynchronous infrared brightness temperatures. In contrast, the morphing algorithm used in IMERG6 is modified to derive cloud motion vectors from variables in the Modern-Era Retrospective Analysis for Research and Applications Version 2 (MERRA-2) reanalysis [27]. Tan and Huffman [28] examined the global precipitation for August 2017 to October 2017 from IMERG5 and IMERG6, and noticed that IMERG6 outperforms IMERG5. However, the possibility that IMERG6 outperforms IMERG5 or any other SPPs with regards to precipitation over Taiwan has not been examined by Tan and Huffman [28].
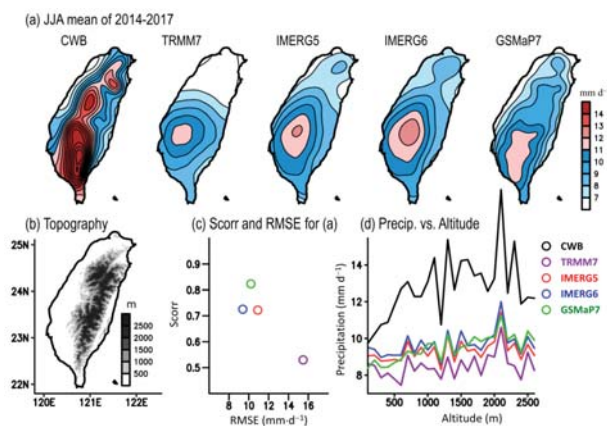
Figure 1. (**a**) Spatial distribution of summer mean precipitation over Taiwan, averaged during the summers (June, July, and August; JJA) of 2014–2017; from left to right is estimations made using Central Weather Bureau (CWB) data, Tropical Rainfall Measuring Mission-3B42 v7 (TRMM7), Integrated Multi-satellitE Retrievals for Global Precipitation Measurement-Final v5 (IMERG5) and v6 (IMERG6), and Global Satellite Mapping of Precipitation for Global Precipitation Measurement-Gauge v7 (GSMaP7). (**b**) The geographic location and topography of Taiwan. (**c**) The spatial correlation (Scorr) and the root mean square error (RMSE) for the comparison between the satellite precipitation products (SPPs) and the CWB data in (**a**). Here, the sample size is 392 grid points for the land areas in (**a**). (**d**) The estimations of the precipitation in (**a**), averaged by area at different altitudes. The color legends of (**c**,**d**) are given in the right panel of (**d**).

The GSMaP project, which was sponsored by Japan Science and Technology Agency during 2002–2007 and extended by JAXA, aims to develop microwave radiometer algorithms for producing high resolution global precipitation maps [29]. After the GPM mission was launched, a new algorithm was developed for the GSMaP project which included the GPM satellite data, producing GSMaP-Gauge data from March 2014 [4]. Recently, Derin et al. [15] pointed out that GSMaP-Gauge v7 (hereafter GSMaP7) and IMERG5 performed better than IMERG6 in depicting the precipitation formation over multiple complex terrain regions, including western Taiwan, during the period 2014–2015. However, only 34 gauges in western Taiwan and only two years of data from 2014–2015 were used by Derin et al. [15] as the reference base for comparison. It should be noted that there are more than 400 rain gauges across the entirety of Taiwan [24] that can be used for a more detailed comparison of SPPs starting from 2000. As the performance of SPPs might be location dependent and timing dependent [24], it is important to examine the performance of SPPs over Taiwan using higher density of rain gauges and longer time periods.

The main objective of this study was to evaluate the performance of multiple SPPs (including TRMM7, IMERG5, IMERG6, and GSMaP7) in depicting the spatial-temporal variations of summer (June, July, and August; JJA) precipitation over whole Taiwan, using more than 400 local rain gauges as the reference base for comparison. The selection of GPM SPPs followed Derin et al. [15]. However, in contrast to Derin et al. [15], who only performed the evaluation at daily and annual timescales, we perform the evaluation of summer precipitation at mean status, daily, interannual and diurnal timescales. In addition, we examine the capabilities to apply SPPs in studying the activities of summer connective afternoon rainfall (CAR) event (Figure 7, explained later), which is the most frequently observed weather pattern in Taiwan [30]. The analysis mainly focuses on the time periods that overlap in all data investigated, that is, the summers of 2014–2017 (Table 1), with an additional comparison between TRMM7 and IMERG6 for the summers of 2000–2017.

Table 1. Information about the satellite precipitation products (SPPs) used in this study.

	TRMM7	IMERG5	IMERG6	GSMaP7
Spatial Resolution	$0.25° \times 0.25°$	$0.1° \times 0.1°$	$0.1° \times 0.1°$	$0.1° \times 0.1°$
Temporal Resolution	3 Hour	Half Hour	Half Hour	Hour
Coverage	50°S–50°N	60°S–60°N	60°S–60°N	60°S–60°N
Period	1998/1–2019/10	2014/3–2018/6	2000/6–present	2014/3–present

The remainder of this manuscript is arranged as follows. Information about the data and the statistical methodology are introduced in Section 2. Section 3 documents the evaluation and application of SPPs in studying the multiple timescale variations of summer precipitation over Taiwan. Discussions are provided in Section 4. A summary is given in Section 5.

2. Data and Methods

2.1. Data

Following Huang et al. [24], the gridded hourly precipitation data produced from 436 rain-gauge observations provided by the Central Weather Bureau (hereafter CWB data) in Taiwan was used as the reference base for comparison. The Cressman scheme [31] was used to generate the gridded CWB data, following the procedures described by Hong and Cao [32].

Table 1 documents the basic information about SPPs used in this study. To compare with hourly CWB data, the 3-hourly TRMM7 were linearly interoperated into the hourly precipitation with a spatial resolution of $0.1° \times 0.1°$. Also, the two half-hourly IMERG estimations (unit: mm·h^{-1}) were averaged to obtain hourly data [24]. All the hourly data were then converted into the local timescale in Taiwan, that is, universal time (UTC) + 8 hours.

The algorithms used by the four SPPs are briefly summarized below. According to Huffman et al. [1], four stages are implemented for the production of TRMM7: (1) precipitation estimates

from the microwave sensor are calibrated and combined; (2) the infrared precipitation estimates are produced to extend the spatial coverage that the microwave observations do not cover; (3) the microwave and infrared precipitation estimates are combined; and (4) the monthly Global Precipitation Climatology Center (GPCC) gauge analysis product is used to climatologically adjust TRMM7 [1,5]. No gauge adjustment is required over the oceans. The TRMM7 data can be obtained from https://pmm.nasa.gov/data-access/downloads/trmm.

According to Huffman et al. [33], the precipitation estimates of IMERG5 are produced using the following steps: (1) the individual satellite sensor data are gridded and calibrated to the combined microwave-radar estimates; (2) the precipitation estimates are propagated forwards and backwards in time using cloud motion vectors derived from infrared data; (3) the propagated precipitation estimates, along with the infrared estimates, are merged based on Kalman weighting factors; and (4) a bias correction is conducted using the monthly GPCC gauge analysis product. Unlike IMERG5, the cloud motion vectors of IMERG6 are derived from MERRA-2 variables [3]. The data of IMERG5 and IMERG6 are available at https://pmm.nasa.gov/data-access/downloads/gpm.

According to Mega et al. [4], several steps are followed for the production of GSMaP7: (1) a simplified and near real-time version of precipitation estimation is generated using fewer passive microwave input streams and a forward-only cloud advection scheme [34]; (2) an improved version of precipitation estimation is generated by applying the Kalman filter to assimilate and refine the visible/infrared-based precipitation rates [29]; (3) both forwards and backwards morphing is applied on the improved version of precipitation estimation to the area observed by the passive microwave radiometer to be affected by precipitation; and (4) a bias correction is conducted using the CPC unified gauge-based analysis of daily precipitation. The GSMaP7 data can be downloaded from https://sharaku.eorc.jaxa.jp/GSMaP/.

2.2. Statistical Methods Applied for Comparison

In this study, the root mean square error (RMSE) for the comparison between the SPPs and the CWB data is calculated based on Equation (1):

$$RMSE \; = \; \sqrt{\sum (SPPs - CWB\ data)^2 / (N-1)}, \tag{1}$$

where N is the sample size [35]. The temporal correlation (Tcorr) [35] and the spatial correlation (Scorr) [35] between the SPPs and the CWB data are also calculated to evaluate the performance of SPPs.

Additionally, following the procedures in earlier literature [36–39], the frequently used threat score (TS) and bias score (BS) are adopted for quantitative evaluation of the precipitation estimations in Taiwan. The values of TS and BS are calculated based on Equations (2) and (3), respectively [36]:

$$TS = H/(O+F-H), \tag{2}$$

$$BS = F/O, \tag{3}$$

where O is the area (i.e., number of grid points) of precipitation depicted by the CWB data that exceeds a given precipitation threshold, F is the area of precipitation depicted by the selected SPP that exceeds the given precipitation threshold, and H is the intersection of O and F over a period of accumulation. The worst and best possible values for TS are 0 and 1, respectively. BS can be described by any value from 0 to infinity. As stated in Levizzani et al. [40], BS gives the ratio of the estimated rain area (frequency) to the observed rain area (frequency), regardless of how well the rain patterns correspond with each other. TS measures the fraction of all events estimated and/or observed that were correctly diagnosed. For other details of TS and BS, please refer to Levizzani [40].

Moreover, to clarify the spatial-temporal characteristics of diurnal precipitation over Taiwan, we applied the widely used empirical orthogonal function (EOF) analysis [41] on the variation of diurnal precipitation. For more details of EOF analysis, please refer to Hannachi et al. [41].

3. Results

3.1. Mean Status

Figure 1a shows the spatial distribution of precipitation over Taiwan, averaged from summers of 2014–2017. From the CWB data, it is noted that larger precipitation is observed in southwest Taiwan than in the other sub-regions. This is due to the interaction between the prevailing southwesterly summer monsoonal flow (not shown) and the local topography (Figure 1b), which can lead to more precipitation occurring on the windward side of the mountains (i.e., southwest Taiwan) [30]. Visually, all four of the SPPs can qualitatively depict the feature with more significant precipitation over southwest Taiwan. However, GSMap7 seems better than the others in illustrating the location of maximum precipitation. This might due to the inclusion of an orographic effect for additional upward motion and moisture flux convergence in the GSMaP algorithm [29,42].

Also worth noting in Figure 1a, all four of the SPPs tend to underestimate the amount of summer mean precipitation over most areas of Taiwan. This might be because that satellite methods underestimate heavy precipitation associated with shallow orographic precipitation systems [42]. Despite the weakness in validating the magnitude of precipitation, a further comparison between the values of precipitation area-averaged over Taiwan indicates that IMERG6 ($\sim$8.9 mm$\cdot$d^{-1}), relative to the other SPPs (TRMM7$\sim$8.0, IMERG5$\sim$8.5, and GSMaP7$\sim$8.6 mm$\cdot$d^{-1}), is closer to the CWB data ($\sim$11.0 mm$\cdot$d^{-1}).

Based on Figure 1a, the Scorr and the RMSE are then calculated for the comparison between the SPPs and the CWB data. As seen in Figure 1c, TRMM7 has the largest RMSE and the smallest Scorr. Compared to the performance of TRMM7 (i.e., Scorr = 0.53, RMSE = 15.5 mm$\cdot$d^{-1}), the Scorr of IMERG6 ($\sim$0.73) is increased by approximately 37.7% [= (0.73 − 0.53)/0.53 × 100%] and the RMSE of IMERG6 ($\sim$9.4 mm$\cdot$d^{-1}) is reduced by approximately 39.4% [= (15.5 − 9.4)/15.5 × 100%]. This suggests that using IMERG6 to replace TRMM7 can increase the performance by approximately 35%~40% in depicting the spatial distribution of summer mean precipitation over Taiwan.

In addition to IMERG6, IMERG5 (i.e., Scorr = 0.72, RMSE = 10.9 mm$\cdot$d^{-1}) and GSMaP7 (i.e., Scorr = 0.82, RMSE = 10.2 mm$\cdot$d^{-1}) also outperform TRMM7 in illustrating the spatial distribution of summer mean precipitation over Taiwan (Figure 1c). Part of the reason for this might be that the original spatial resolution of IMERG5, IMERG6, and GSMaP7 (0.1° × 0.1°) is higher than that of TRMM7 (0.25° × 0.25°). Therefore, although TRMM7 has been re-gridded into the same spatial resolution (0.1° × 0.1°) for the comparison (see Section 2), the performance of TRMM7 is still worse than the other higher resolution SPPs for depicting the precipitation that occurs over complex terrain.

To reveal whether the above suggestion is dependent on the altitude, we further compared the performance of SPPs at different altitudes. Worth noting in Figure 1d, there is an obvious increase in the difference between the SPPs and the CWB data, as the altitude increases. Among the four SPPs, the performance of IMERG6 (GSMaP7) is closer to the CWB data at most altitudes below 1000 m (higher than 1500 m), while TRMM7 has the largest bias over most altitudes. These features indicate again that the new GPM SPPs (including IMERG5, IMERG6, and GSMaP7) outperform TRMM7 in depicting the distribution of summer mean precipitation over Taiwan; this finding is not dependent on the altitude.

3.2. Day-to-Day and Interannual Variation

Figure 2a shows the time series of the 5-day running mean for the precipitation area-averaged over Taiwan during the summers of 2014–2017. Using the time series in Figure 2a, two statistical scores, Tcorr and RMSE, between the CWB data and the SPPs, were then calculated for providing evaluation evidence. As seen in Figure 2b, all SPPs have similar values of Tcorr ($\sim$0.9), but the lowest RMSE is observed in IMERG6 and the highest RMSE is observed in TRMM7. Overall, the RMSE of IMERG6 ($\sim$7.6 mm$\cdot$d^{-1}) is approximately 12.6% lower [= (8.7 − 7.6)/8.7 × 100%] than the RMSE of TRMM7 ($\sim$8.7 mm$\cdot$d^{-1}).
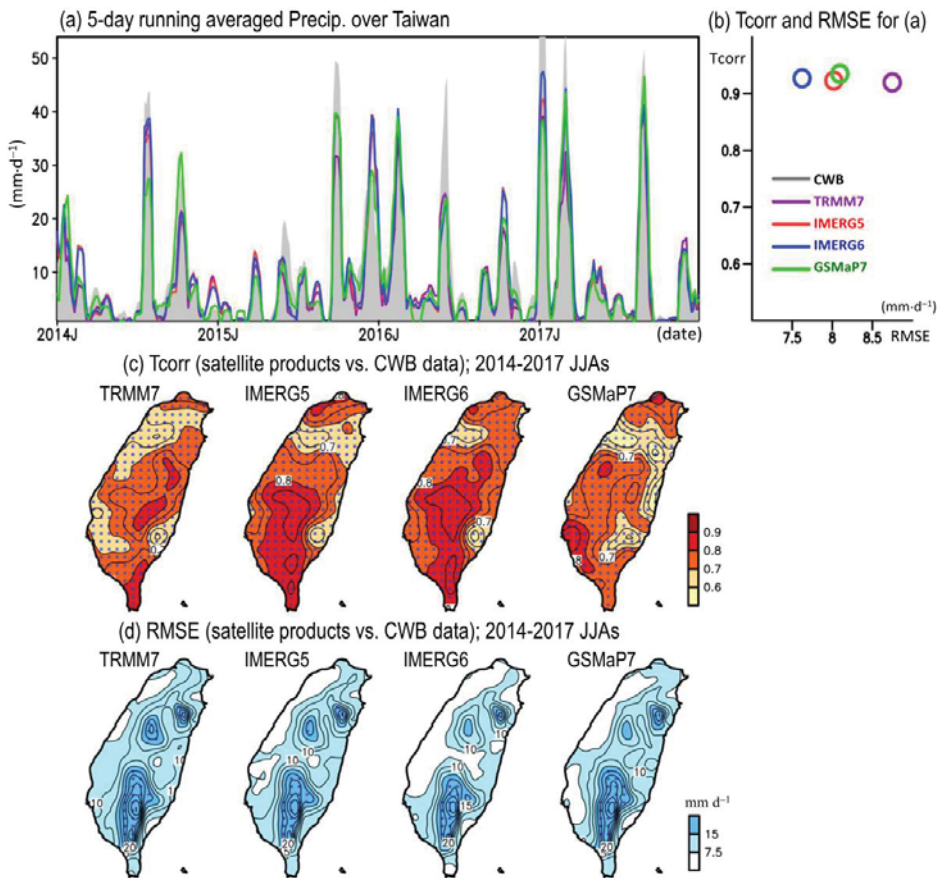
Figure 2. (**a**) Time series of 5-day running averaged precipitation over Taiwan, during the summers of 2014–2017, extracted from the CWB data and the four SPPs. The color symbols are described in (**b**). (**b**) Tcorr (i.e., temporal correlation) and RMSE between the time series of CWB data and SPPs in (**a**). (**c**) Grid-to-grid Tcorr between the time series of CWB data and SPPs. (**d**) Grid-to-grid RMSE between the time series of CWB data and SPPs. Here, the sample size for calculating Tcorr and RMSE is 368 days from summers of 2014–2017. In (**c**,**d**), the values pass the 99% significant test are marked by dots.

Additionally, as noted from Figure 2c, which shows the spatial distribution of grid-to-grid Tcorr for the comparison between the SPPs and the CWB data, IMERG6 has more areas with larger values of Tcorr (e.g., >0.8). Moreover, even though the related spatial distributions of grid-to-grid RMSE in Figure 2d do not show too much difference among the performance of the four SPPs, IMERG6 is still the one having more areas with smaller values of RMSE (e.g., <7.5 mm·d^{-1}). These features suggest that the performance of IMERG6 is overall better than the other SPPs in depicting the day-to-day variations of precipitation over Taiwan.

Recall, Derin et al. [15] indicated that IMERG6 performed worse than GSMaP7 and IMERG5 in capturing the daily precipitation formation over western Taiwan during 2014–2015. Consistent with Derin et al. [15], one can note from Figure 2c that GSMaP7 and IMERG5 did perform better than IMERG6 in some coastal regions of southwest Taiwan, even though the time periods used for the comparison are different in Figure 2c and Derin et al. [15]. However, in contrast to Derin et al. [15], we would like to call attention that when focused on the daily precipitation formation over whole

Taiwan during the summers of 2014–2017, the performance of IMERG6 is overall better than GSMaP7 and IMERG5.

Next, statistical evidence is provided for evaluating the capabilities of IMERG6 and TRMM7 to depict the variations of daily precipitation events during the summers of 2000–2017. Figure 3a shows the distribution of the occurrence frequency of precipitation events at various ranges of intensity (units: $mm \cdot d^{-1}$). From Figure 3a, we note that both TRMM7 and IMERG6 tend to underestimate the occurrence frequency of precipitation events at most ranges of intensity. Despite that, the performance of IMERG6 is overall better (i.e., more close to the CWB data) than TRMM7, in particularly for capturing the occurrence frequency of light precipitation events (see right top panel of Figure 3a).
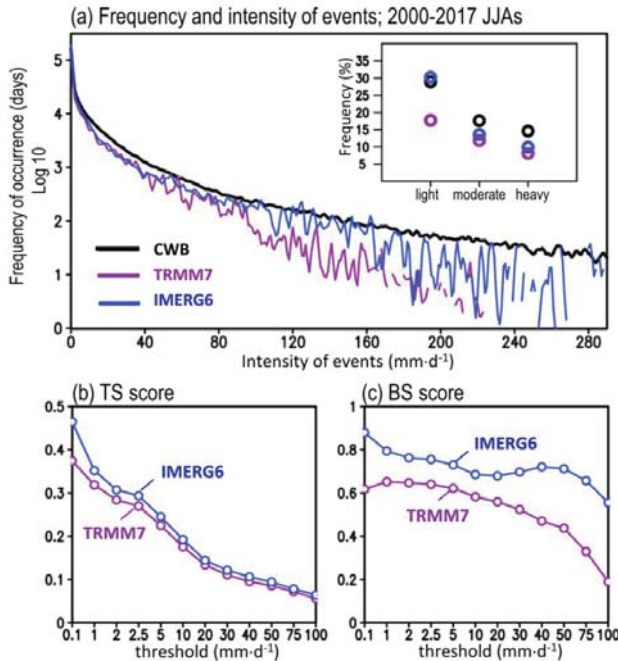
Figure 3. (**a**) Histograms of the frequency of occurrence as a function of daily precipitation intensity (bin size is 1 $mm \cdot d^{-1}$) in Taiwan during the summers of 2000–2017. Inset plots represent the frequency (in %) of light (0.1–5 $mm \cdot d^{-1}$), moderate (5–20 $mm \cdot d^{-1}$), and heavy precipitation (>20 $mm \cdot d^{-1}$) events. The method used for generating (**a**) follows Sun et al. [43]. (**b**) and (**c**) is the value of TS and BS (explained in Section 2), respectively, for the comparison between CWB data and two SPPs (IMERG6 and TRMM7) during the summers of 2000–2017. The sample size used here is 92 (days per JJA) × 18 (JJAs) × 392 (grid points per day) = 649152 grid points.

Moreover, two other statistical scores, TS (i.e., threat score) and BS (i.e., bias score), are calculated for representing the skill of SPPs in quantitative precipitation estimations. It can be noted in Figure 3b that the value of TS in IMERG6 is higher than in TRMM7 over all ranges of precipitation threshold. As higher TS values indicate better performance [36], Figure 3b again suggests that IMERG6 outperforms TRMM7. On the other hand, both IMERG6 and TRMM7 have values of BS < 1 over all ranges of precipitation threshold. This implies that precipitation events in both IMERG6 and TRMM7 occurred less often than that in the CWB data [36]. However, relative to TRMM7, IMERG6 still has a BS value closer to the ideal value of 1.

Figure 4a shows the interannual variation of summer precipitation, area-averaged over Taiwan, estimated by the CWB data and the SPPs. Consistent with Figure 3a, the variations of IMERG6

(TRMM7) is more (less) close to the CWB data during the summers of 2000–2017. The spatial distribution of grid-to-grid Tcorr (RMSE) between the SPPs and the CWB data are further conducted in Figure 4b (Figure 4c) for evaluating the performance of TRMM7 and IMERG6. It is noted that IMERG6 (TRMM7) has more (less) areas with Tcorr > 0.8 and RMSE < 5 mm·d^{-1}, suggesting again that IMERG6 outperforms TRMM7.
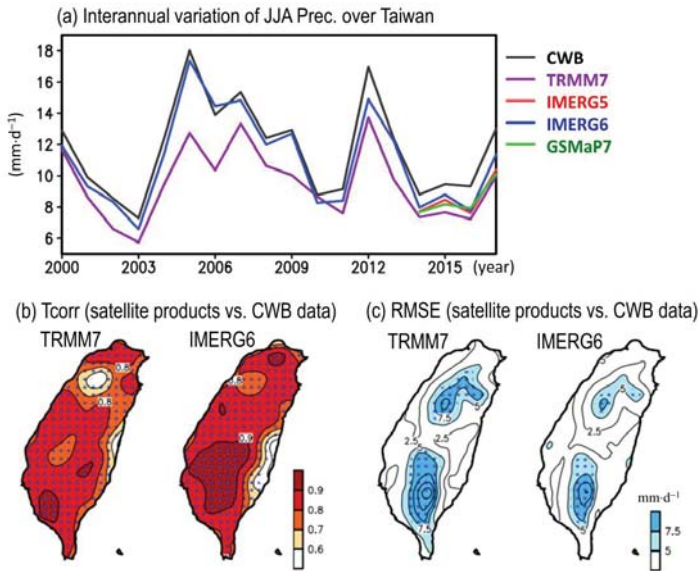
Figure 4. (**a**) Time series of the mean precipitation for the summers of 2000–2017, area-averaged over Taiwan, as estimated by the CWB data and SPPs. (**b**) and (**c**) is the related grid-to-grid Tcorr and RMSE, respectively, between the CWB data and the selected SPPs: TRMM7 and IMERG6. Here, the sample size for calculating Tcorr and RMSE is 18 JJAs. In (**b**,**c**), the values pass the 99% significant test are marked by dots.

Also, for the performance of IMERG5 and GSMaP7 in illustrating the interannual variation of summer precipitation over Taiwan, some information is given in Figure 4a. It was found that even when focused on the summers of 2014–2017, IMERG6 (TRMM7) is still more (less) close to the CWB data, as compared to the other SPPs.

3.3. Diurnal Variation

The summer precipitation over Taiwan also exhibits a clear diurnal feature in which the maximum precipitation generally occurs in the afternoon [24,30]. To illustrate this feature, we apply an EOF analysis on the variation of diurnal precipitation, averaged during the summers of 2014–2017. This analysis method is frequently adopted by earlier studies in examining the characteristics of diurnal variation of precipitation over East Asia [44–46]. Here, only the first mode of the EOF analysis is presented in Figure 5, and several features noted from Figure 5 are summarized below:

1. The first EOF mode of the CWB data explains about 62.6% of the total variability of diurnal precipitation. Among the four SPPs, IMERG6 (GSMaP7) has the number of percentage more (less) close to the CWB data.
2. Spatially (Figure 5a), the first EOF mode of the CWB data shows two maximum centers: one in central-west Taiwan and the other in northern Taiwan. For the SPPs, all of them are able to show

two maximum centers in Figure 5a. However, the locations of the maximum centers in GSMaP7 are apparently shifted to the west compared to the CWB data.

3. Temporally (Figure 5b), the first EOF mode of the CWB data shows the timing of diurnal precipitation maximum occurred between 15–18 h. For the SPPs, all of them are able to capture similar feature in Figure 5b, even though their amplitudes of diurnal variation are weaker than the CWB data. Among the four SPPs, GSMaP7 underestimates the most in the amplitude of diurnal variation of precipitation.
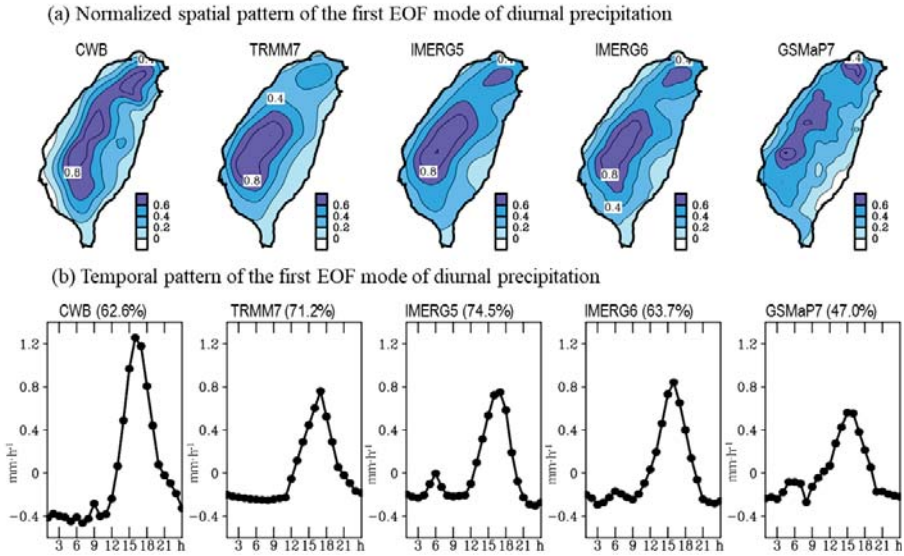
Figure 5. The first empirical orthogonal function (EOF) mode of diurnal precipitation, averaged during the summers of 2014–2017: (**a**) the normalized spatial patterns (i.e., eigen-vectors normalized to the maximum precipitation), and (**b**) the temporal patterns (i.e., eigen-coefficients). The percentages (%) of the total variability in hourly precipitation explained by the first EOF mode are added in (**b**).

Based on Figure 5, it seems that GSMaP7 performs worse than the other SPPs in illustrating the diurnal variation of precipitation. Indeed, by calculating the Scorr (RMSE) between the SPPs and the CWB data from Figure 5a, we note from Table 2 that GSMaP7 has the lowest (highest) value of Scorr (RMSE), suggesting its spatial pattern is less similar to the CWB data. In addition, by calculating the Tcorr (RMSE) between the SPPs and the CWB data from Figure 5b, we note from Table 2 that GSMaP7 has the lowest (highest) value of Tcorr (RMSE), suggesting its temporal pattern is also less similar to the CWB data.

It is also apparent from Table 2 that IMERG6 has the smallest bias in capturing the spatial-temporal characteristics of the diurnal precipitation over Taiwan. Tan et al. [47] examined the performance of IMERG6 in capturing the diurnal cycle of precipitation over the southeastern United States also noted that IMERG6 tends to underestimate the diurnal amplitude, but is capable of depicting the phase of diurnal precipitation. However, why IMERG6 (GSMaP7) performs the best (worst) with regards to the diurnal precipitation over Taiwan is unclear and requires further study.

Table 2. Statistical values for the comparison between the CWB data and SPPs shown in Figure 5. The unit of RMSE is mm·d^{-1}. The lowest value of Scorr and Tcorr, as well as the highest value of RMSE, are marked by *. The sample size of spatial variation is 392 grid points. The sample size of temporal variation is 24 hours.

	TRMM7	IMERG5	IMERG6	GSMaP7
Scorr for Figure 5a	0.83	0.88	0.89	0.74*
RMSE for Figure 5a	0.17	0.16	0.14	0.20*
Tcorr for Figure 5b	0.92	0.94	0.94	0.85*
RMSE for Figure 5b	0.26	0.24	0.23	0.35*

The performance of IMERG6 and TRMM7 in illustrating the variation of diurnal precipitation area-averaged over Taiwan during the summers of 2000–2017 was further evaluated based on Figure 6. The CWB data (Figure 6a) shows that all examined periods have maximum diurnal precipitation occurred between 15-18 h. By comparing Figure 6a with Figure 6b,c, we note that both TRMM7 and IMERG6 are able to show the temporal phase evolution similar to the CWB data, with the value of Scorr between Figures 6a and 6b (Figure 6c) is about 0.91 (0.92). However, it is also apparent in Figure 6 that both TRMM7 and IMERG6 tend to underestimate the amplitude of diurnal precipitation for all examined time periods, but TRMM7 (IMERG6) is less (more) close to the CWB data. All above features revealed in Figure 6 are consistent with those suggested by Figure 5 and Table 2, suggesting again that IMERG6 is better than TRMM7 in depicting the variation of diurnal precipitation over Taiwan.
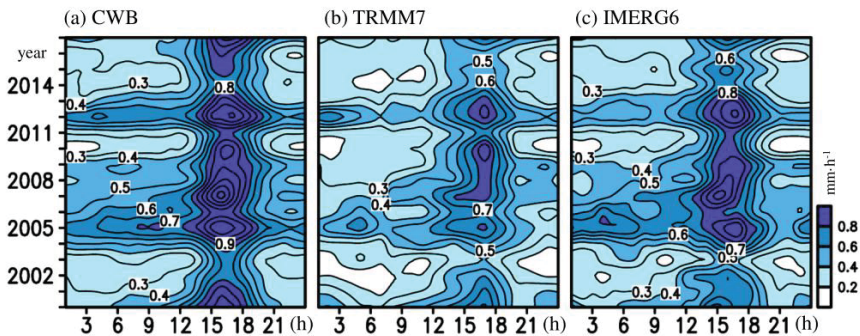
Figure 6. Temporal evolution of the hourly precipitation area-averaged over Taiwan extracted from the selected data for each specific summer during 2000–2017: (**a**) the CWB data, (**b**) TRMM7, and (**c**) IMERG6.

3.4. Potential Applications

Based on findings of Sections 3.1–3.3, we then infer that applying IMERG6 to the study of summer convective afternoon rainfall (CAR) events over Taiwan (e.g., Figure 7a), which generally includes a diurnal precipitation maximum in the afternoon after the local thermal heating maximum (e.g., Figure 7b), can obtain results more similar to those seen in the CWB data. This inference will be clarified by the examinations presented in this sub-section. Hereafter, the methods used for the identification of CAR events follow Huang et al. [30], and are briefly summarized as follows: (1) a rainy day is defined as a day with an accumulated precipitation of ≥ 0.1 mm; (2) the accumulated precipitation of a rainy day during the time period 1200–2200 h is > 80% of the daily precipitation; (3) the accumulated precipitation of a rainy day during the time period 0100–1100 h is < 10% of the daily precipitation; and (4) days affected by other weather systems (e.g., typhoons and frontal systems) are excluded.
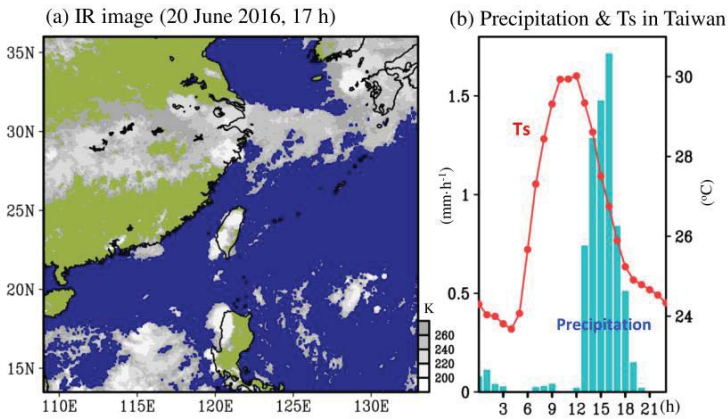
Figure 7. (**a**) Infrared cloud image, obtained from Gridded Satellite B1 Observations (https://www.ncdc.noaa.gov/gridsat/), for an example of convective afternoon rainfall (CAR) event that occurred on 20 June 2016, 17 h (local time) in Taiwan. (**b**) Time series of hourly precipitation (bars) and surface temperature (Ts, red line) averaged from local stations in Taiwan for the event shown in (**a**).

Figure 8 shows the spatial distribution for the contribution of CAR activities (including occurrence frequency and precipitation amount) to the total summer precipitation events, averaged over the summers of 2014–2017. In Figure 8a (Figure 8b), the CWB data shows that CAR events contribute more than 40% (30%) of the occurrence frequency (precipitation amount) of the total precipitation events in most areas of Taiwan. Furthermore, similar to Figure 5a, two maximum centers are revealed in the CAR activities observed by the CWB data. Despite the location difference, all SPPs are able to show two maximum centers in the CAR activities; however, GSMaP7 apparently underestimates the most in the contribution of CAR activities to the total precipitation events.
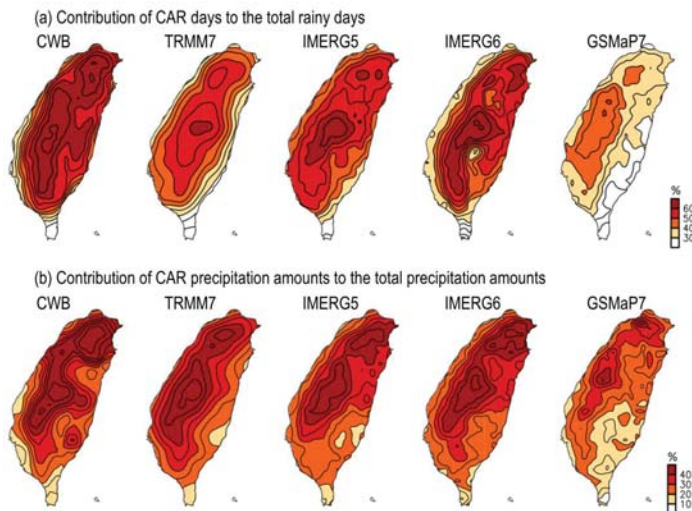
Figure 8. Contribution of CAR events to total precipitation events during the summers of 2014–2017: (**a**) frequency of occurrence, (**b**) amount of precipitation. In (**a**,**b**), contribution = $\frac{\text{CAR events}}{\text{Total precipitation events}} \times 100\%$.

By comparing the CWB data with the four SPPs in Figure 8, two statistical scores (Scorr and RMSE) were calculated and documented in Table 3. It can be noted in Table 3 that IMERG6 (with the highest Scorr and the lowest RMSE) outperforms the other SPPs, while the greatest bias is seen in GSMaP7. This is also consistent with what revealed in Table 2, suggesting that the higher (lower) performance of IMERG6 (GSMaP7) in depicting CAR activities may be attributed to its higher (lower) performance in illustrating the diurnal variation of precipitation over Taiwan.

Table 3. Statistical values for the comparison between the CWB data and SPPs shown in Figure 8. The unit of RMSE is %. The lowest value of Scorr and the highest value of RMSE are marked by *. The sample size is 392 grid points.

	TRMM7	IMERG5	IMERG6	GSMaP7
Scorr for Figure 8a	0.78	0.82	0.84	0.54*
RMSE for Figure 8a	12.08	8.82	8.48	21.08*
Scorr for Figure 8b	0.74	0.79	0.83	0.69*
RMSE for Figure 8b	8.28	7.67	6.92	12.67*

Additionally, we examine the performance of TRMM7 and IMERG6 in depicting the interannual variation of CAR activities area-averaged over Taiwan during the summers of 2000–2017. Figure 9a shows that both TRMM7 and IMERG6 are capable of depicting the interannual variation of contribution of CAR events to the total precipitation amount, similar to those seen in the CWB data. The Tcorr between the time series of TRMM7 (IMERG6) and the CWB data in Figure 9a is approximately 0.73 (0.80), which passes the 99% significant test.
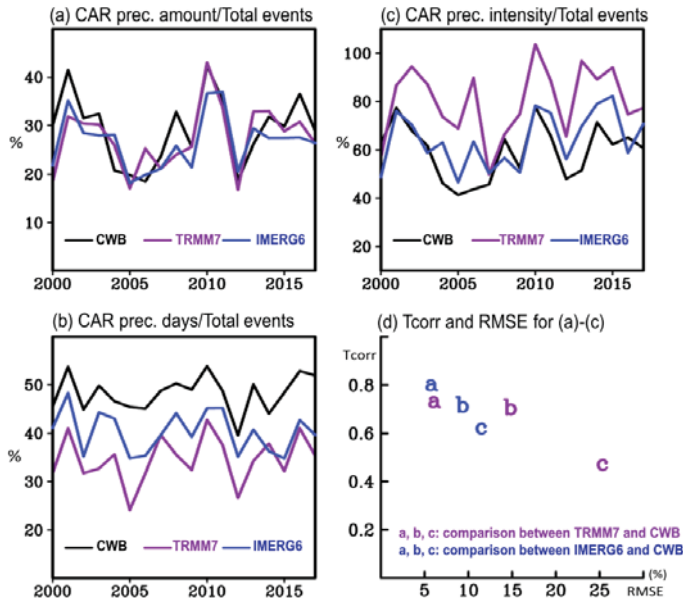
Figure 9. (**a**) The contribution of CAR events to the total precipitation amount that fell during the summers of 2000–2017. (**b**) The contribution of CAR days to total rainy days. (**c**) The contribution of CAR intensity (i.e., mean rain rate; unit: mm per event) to the intensity of total rainy events. In (**a**–**c**), contribution = $\frac{\text{CAR events}}{\text{Total precipitation events}} \times 100\%$. In (**d**), the comparison of Tcorr and RMSE between time series in (**a**–**c**) is given, using the CWB data as the reference base, and denoted a-b-c, respectively. In (**d**), the purple color and blue color represents results related to TRMM7 and IMERG6, respectively.

By separating the precipitation amount into the occurrence frequency and the intensity (i.e., precipitation amount = occurrence frequency × intensity), it is however noted from Figure 9b that both IMERG6 and TRMM7 tend to underestimate the contribution of CAR events to the occurrence frequency of total precipitation events, as compared to the CWB data. This might be because the SPPs, which utilize precipitation estimation from infrared and passive microwave sensors, are poor at retrieving local precipitation events over complex mountainous areas [42,48].

In contrast to Figure 9b, it can be noted in Figure 9c that both IMERG6 and TRMM7 tend to overestimate the contribution of CAR events to the intensity of total precipitation events. This might be because that satellite methods assume heavy precipitation results from deep clouds [42], and CAR events in Taiwan are belong to local deep convections [30]. Despite the bias seen in Figure 9b,c, we note from Figure 9d that IMERG6 (with higher Tcorr and lower RMSE) outperforms TRMM7 overall in illustrating the interannual variations of CAR activities. Therefore, we suggest that using IMERG6 to replace TRMM7 can benefit the researcher by obtaining more accurate characteristics of CAR in Taiwan.

4. Discussions

It was noted in Figure 1a that more (less) differences are observed between GSMaP7 (IMERG5 and IMERG6) and TRMM7 concerning the distribution of maximum precipitation over Taiwan. This might be because the land precipitation in GSMaP7 is adjusted using the CPC gauge-based analysis daily precipitation, while the land precipitation in TRMM7 and the IMERG products are adjusted using the GPCC gauge-based monthly precipitation [1,3,4]. In fact, by comparing the difference between the mean status of GSMaP7 (IMERG5 and IMERG6) and TRMM7 over a larger domain averaged during the summers of 2014–2017, we note that the difference between GSMaP7 and TRMM is also larger over Indochina (see Figure 10). In contrast, over the ocean areas of the Asian domain, larger differences can be found between IMERG5 and TRMM7.
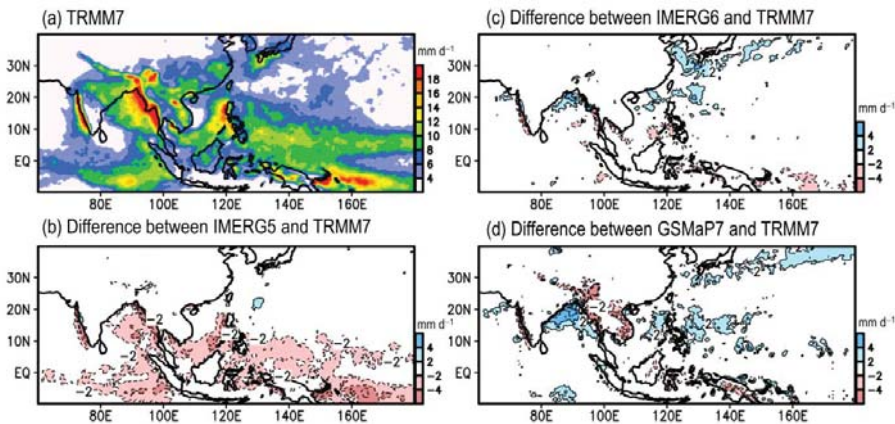
Figure 10. (**a**) Mean seasonal precipitation averaged during the summers of 2014–2017, estimated by TRMM7, over the Asian domain (60°E–180°, 10°S–40°N). (**b**) is related to (**a**), except for the difference between IMERG5 and TRMM7. (**c**) is similar to (**b**), except for the difference between IMERG6 and TRMM7. (**d**) is similar to (**b**), except for the difference between GSMaP7 and TRMM7.

Consistent with Figure 10, IMERG5 also has the daily variations in oceanic precipitation area-averaged over the Asian domain, more different to TRMM7 (Figure 11b; 5 day running mean applied). In contrast, for the daily variations in land-averaged precipitation, GSMaP7 appears to differ the most from TRMM7 (Figure 11a; 5 day running mean applied). To support the above arguments, we further calculate Tcorr and RMSE based on the time series given in Figure 11a,b and show the results in Figure 11c. It is confirmed that: (1) GSMaP7 has a more clear difference to TRMM7 (i.e., smaller

Tcorr and larger RMSE) over the land; and (2) IMERG5 has a more clear difference to TRMM7 over the ocean. This information is particularly important for studies that tend to use GPM products to replace TRMM7 for estimating the moisture budget over the Asian domain [49–52].
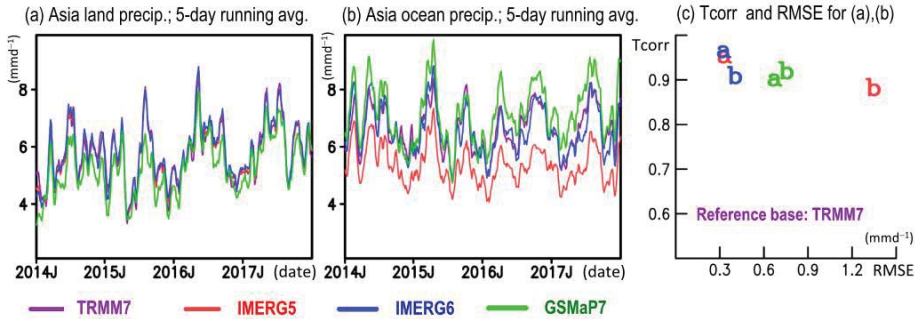
Figure 11. Time series of 5-day running averaged precipitation area-averaged over (**a**) the land and (**b**) the ocean of the Asian domain (60°E–180°, 10°S–40°N). The time series covered is from 2014 to 2017 JJAs (i.e., 368 days). (**c**) is based on (**a**,**b**), except for the calculation of Tcorr and RMSE between TRMM7 and the selected SPPs, including IMERG5 (red), IMERG6 (blue), and GSMaP7 (green).

5. Conclusions

This study evaluated the performance of multiple SPPs (including TRMM7, IMERG5, IMERG6, and GSMaP7) in depicting the variations of summer precipitation over Taiwan. The evaluations were performed at mean status, daily, interannual, and diurnal timescales. The major goal was to clarify whether IMERG6 is the best choice to replace TRMM7 for studying the summer precipitation variations over Taiwan. In addition, capabilities to apply SPPs in studying the most frequently observed weather pattern in Taiwan, that is, CAR event, is also examined. The performances of SPPs were evaluated using more than 400 gauges (i.e., CWB data) in Taiwan as the reference base for comparison.

Our analyses show that IMERG6 overall outperforms the other SPPs for depicting the spatial-temporal characteristics of summer precipitation over Taiwan varied at multiple timescales. Among the four SPPs, IMERG6 also has the smallest bias in depicting the characteristics of CAR activities in Taiwan. All the above findings suggest that more accurate results can be obtained if IMERG6 is used to replace TRMM7 for studying issues that are related to the summer precipitation in Taiwan. It is noted that above statement is true for the time period between 2000 and 2017. Further studies are suggested to evaluate the performance of multiple SPPs in Taiwan during the dry seasons.

Additionally, when investigating the differences between TRMM7 and the other SPPs over the Asian domain, we note that IMERG5 (GSMaP7) has the largest difference to TRMM7 over the ocean (land). This finding provides useful information for future studies that tend to use GPM products to replace TRMM7 for estimating the moisture budget over the Asian domain e.g., [49–52].

References

1. Huffman, G.J.; Bolvin, D.T.; Neklin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.J.; 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]

2. 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. Meteor. Soc.* **2014**, *95*, 701–722. [CrossRef]

3. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Stocker, E.F.; Tan, J. *V06 IMERG Release Notes*; NASA/GSFC: Greenbelt, MD, USA, 2019.

4. Mega, T.; Ushio, T.; Takahiro, M.; Kubota, T.; Kachi, M.; Oki, R. Gauge-Adjusted Global Satellite Mapping of Precipitation. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 1928–1935. [CrossRef]

5. Tian, Y.; Peters-Lidard, C.D.; Choudhury, B.J.; Garcia, M. Multitemporal Analysis of TRMM-Based Satellite Precipitation Products for Land Data Assimilation Applications. *J. Hydrometeorol.* **2007**, *8*, 1165–1183. [CrossRef]

6. 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*, 5. [CrossRef]

7. Sungmin, O.; Foelsche, U.; Kirchengast, G.; Fuchsberger, J.; Tan, J.; Petersen, W.A. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6559–6572. [CrossRef]

8. Lu, D.; Yong, B. Evaluation and Hydrological Utility of the Latest GPM IMERG V5 and GSMaP V7 Precipitation Products over the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 2022. [CrossRef]

9. Wang, C.; Tang, G.; Han, Z.; Guo, X.; Hong, Y. Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements. *J. Hydrol.* **2018**, *564*, 342–356. [CrossRef]

10. Lu, X.Y.; Tang, G.Q.; Wang, X.Q.; Liu, Y.; Ji, L.H.; Xie, G.H.; Li, S.; Zhang, Y.X. Correcting GPM IMERG precipitation data over the Tianshan Mountains in China. *J. Hydrol.* **2019**, *575*, 1239–1252. [CrossRef]

11. Mazzoglio, P.; Laio, F.; Balbo, S.; Boccardo, P.; Disabato, F. Improving an Extreme Rainfall Detection System with GPM IMERG data. *Remote Sens.* **2019**, *11*, 677. [CrossRef]

12. Salles, L.; Frédéric, S.; Roig, H.; Almeida, T.; Olivetti, D.; Ferreira, W. Seasonal effect on spatial and temporal consistency of the new GPM-based IMERG-v5 and GSMaP-v7 satellite precipitation estimates in Brazil's central plateau region. *Water* **2019**, *11*, 668. [CrossRef]

13. 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]

14. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM Precipitation Products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]

15. Derin, Y.; Anagnostou, E.; Berne, A.; Borga, M.; Boudevillain, B.; Buytaert, W.; Chang, C.-H.; Chen, H.; Delrieu, G.; Hsu, Y.C.; et al. Evaluation of GPM-era Global Satellite Precipitation Products over Multiple Complex Terrain Regions. *Remote Sens.* **2019**, *11*, 2936. [CrossRef]

16. 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]

17. Tang, G.; Zeng, Z.; Long, D.; Guo, X.; Yong, B.; Zhang, W.; Hong, Y. Statistical and Hydrological Comparisons between TRMM and GPM Level-3 Products over a Midlatitude Basin: Is Day-1 IMERG a Good Successor for TMPA 3B42V7? *J. Hydrometeorol.* **2016**, *17*, 121–137. [CrossRef]

18. Tan, M.L.; Santo, H. Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia. *Atmos. Res.* **2017**, *202*, 63–76. [CrossRef]

19. Xu, R.; Tian, F.; Yang, L.; Hu, H.; Lu, H.; Hou, A. Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain-gauge network. *J. Geophys. Res. Atmos.* **2017**, *122*, 910–924. [CrossRef]

20. Jiang, S.; Ren, L.; Xu, C.Y.; Yong, B.; Yuan, F.; Liu, Y.; Yang, X.; Zeng, X. Statistical and hydrological evaluation of the latest Integrated Multi-satellitE Retrievals for GPM (IMERG) over a midlatitude humid basin in South China. *Atmos. Res.* **2018**, *214*, 418–429. [CrossRef]

21. Jiang, L.G.; Bauer-Gottwein, P. How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. *J. Hydrol.* **2019**, *572*, 486–500. [CrossRef]

22. Wu, Y.; Zhang, Z.; Huang, Y.; Jin, Q.; Chen, X.; Chang, J. Evaluation of the GPM IMERG v5 and TRMM 3B42 v7 Precipitation Products in the Yangtze River Basin, China. *Water.* **2019**, *10*, 1459. [CrossRef]

23. Chen, S.Y.; Huang, W.R. Evaluation on the performance of TRMM, CMORPH, and PERSIANN in depicting the diurnal precipitation variation in Taiwan. *Atmos. Sci.* **2017**, *45*, 167–191. (In Chinese)

24. Huang, W.R.; Chang, Y.H.; Liu, P.Y. Assessment of IMERG precipitation over Taiwan at multiple timescales. *Atmos. Res.* **2018**, *214*, 239–249. [CrossRef]

25. Kishtawal, C.M.; Krishnamurti, T.N. Diurnal variation of summer rainfall over Taiwan and its detection using TRMM observations. *J. Appl. Meteor. Climatol.* **2001**, *40*, 331–344. [CrossRef]

26. Yeh, N.C.; Chuang, Y.C.; Peng, H.S.; Hsu, K.L. Bias Adjustment of Satellite Precipitation Estimation Using Ground-Based Observation: Mei-Yu Front Case Studies in Taiwan. *Asia-Pac. J. Atmos. Sci.* **2019**. [CrossRef]

27. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.* **2017**, *30*, 5419–5454. [CrossRef]

28. Tan, J.; Huffman, G.J. *Computing Morphing Vectors for Version 06 IMERG*; NASA/GSFC: Greenbelt, MD, USA, 2019.

29. Ushio, T.; Sasashige, K.; Kubota, T.; Shige, S.; Okamoto, K.; Aonashi, K.; Inoue, T.; Takahashi, N.; Iguchi, T.; Kachi, M.; et al. A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteorol. Soc. Jpn.* **2009**, *87A*, 137–151. [CrossRef]

30. Huang, W.R.; Chang, Y.H.; Cheng, C.T.; Hsu, H.H.; Tu, C.Y.; Kitoh, A. Summer Convective Afternoon Rainfall Simulation and Projection using WRF Driven by Global Climate Model. Part I: Over Taiwan. *Terr. Atmos. Ocean.* **2016**, *27*, 659–671. [CrossRef]

31. Cressman, G.P. An operational objective analysis system. *Mon. Weather Rev.* **1959**, *87*, 367–374. [CrossRef]

32. Hong, J.S.; Cao, J.H. Objective analysis of the rain gauge observations by using the Cressman analysis method. *Atmos. Sci.* **2011**, *39*, 201–213. (In Chinese)

33. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Kidd, C.; Nelkin, E.J.; Sorooshian, S.; Tan, J.; Xie, P. *Algorithm Theoretical Basis Document (ATBD) Version 5.1: NASA Global Precipitation Measurement (GPM) Integrated Multi-SatellitE Retrievals for GPM (IMERG)*; NASA/GSFC: Greenbelt, MD, USA, 2017.

34. Kubota, T.; Shige, S.; Hashizume, H.; Ushio, T.; Aonashi, K.; Kachi, M.; Okamoto, K. Global precipitation map using satellite-borne microwave radiometers by the GSMaP Project: Production and validation. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2259–2275. [CrossRef]

35. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 1st ed.; Academic Press: Cambridge, MA, USA, 1995; p. 467.

36. Wang, C. On the Calculation and Correction of Equitable Threat Score for Model Quantitative Precipitation Forecasts for Small Verification Areas: The Example of Taiwan. *Weather Forecast.* **2014**, *29*, 788–798. [CrossRef]

37. Wang, C.; Kung, C.; Lee, C.; Chen, T.J. Development and Evaluation of Mei-Yu Season Quantitative Precipitation Forecasts in Taiwan River Basins Based on a Conceptual Climatology Model. *Weather Forecast.* **2012**, *27*, 586–607. [CrossRef]

38. Wang, C.; Huang, S.; Chen, S.; Chang, C.; Tsuboki, K. Cloud-Resolving Typhoon Rainfall Ensemble Forecasts for Taiwan with Large Domain and Extended Range through Time-Lagged Approach. *Weather Forecast.* **2016**, *31*, 151–172. [CrossRef]

39. Yang, T.H.; Yang, S.C.; Ho, J.Y.; Lin, G.F.; Hwang, G.D.; Lee, C.S. Flash flood warnings using the ensemble precipitation forecasting technique: A case study on forecasting floods in Taiwan caused by typhoons. *J. Hydrol.* **2015**, *520*, 367–378. [CrossRef]

40. Levizzani, V.; Bauer, P.; Turk, F.J. *Measuring Precipitation from Space: EURAINSAT and the Future*; Springer Netherlands: Heidelberg, Germany, 2007; p. 724.

41. Hannachi, A.; Jolliffe, I.; Stephenson, D. Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.* **2007**, *27*, 1119–1152. [CrossRef]

42. Shige, S.; Kida, S.; Ashiwake, H.; Kubota, T.; Aonashi, K. Improvement of TMI rain retrievals in mountainous areas. *J. Appl. Meteor. Climatol.* **2013**, *52*, 242–254. [CrossRef]

43. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K.L. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.* **2018**, *56*, 79–107. [CrossRef]

44. Teo, C.-K.; Koh, T.-Y.; Lo, J.; Bhatt, B. Principal Component Analysis of Observed and Modeled Diurnal Rainfall in the Maritime Continent. *J. Clim.* **2011**, *24*, 4662–4675. [CrossRef]

45. Huang, W.R.; Wang, S.S.-Y. Future changes in propagating and non-propagating diurnal rainfall over East Asia. *Clim. Dyn.* **2017**, *49*, 375–389. [CrossRef]

46. Huang, W.R.; Chang, Y.H. Characteristics and mechanisms of the diurnal variation of winter precipitation in Taiwan. *Int. J. Climatol.* **2018**, *38*, 3058–3068. [CrossRef]

47. Tan, J.; Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Diurnal Cycle of IMERG V06 Precipitation. *Geophys. Res. Lett.* **2019**, *46*, 13584–13592. [CrossRef]

48. Derin, Y.; Yilmaz, K.K. Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. *J. Hydrometeor.* **2014**, *15*, 1498–1516. [CrossRef]

49. Wang, Z.; Duan, A.; Yang, S.; Ullah, K. Atmospheric moisture budget and its regulation on the variability of summer precipitation over the Tibetan Plateau. *J. Geophys. Res. Atmos.* **2016**, *122*, 614–630. [CrossRef]

50. Chen, J.; Bordoni, S. Intermodel spread of East Asian summer monsoon simulations in CMIP5. *Geophys. Res. Lett.* **2014**, *41*, 1314–1321. [CrossRef]

51. Prasanna, V.; Annamalai, H. Moist dynamics of extended monsoon breaks over South Asia. *J. Clim.* **2012**, *25*, 3810–3831. [CrossRef]

52. Unnikrishnan, C.K.; Rajeevan, M. Atmospheric water budget over the South Asian summer monsoon region. *Meteorol. Atmos. Phys.* **2018**, *130*, 175–190. [CrossRef]

![remote sensing logo]

![MDPI logo]

Article

Seasonal Characteristics of Disdrometer-Observed Raindrop Size Distributions and Their Applications on Radar Calibration and Erosion Mechanism in a Semi-Arid Area of China

Zongxu Xie [1,2], Hanbo Yang [1,2,*], Huafang Lv [1,2] and Qingfang Hu [3]

[1] Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China;
xie-zx19@mails.tsinghua.edu.cn (Z.X.); lvhf@tsinghua.edu.cn (H.L.)

[2] State Key Laboratory of Hydro-Science and Engineering, Tsinghua University, Beijing 100084, China

[3] State Key Laboratory of Hydro-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China; qfhu@nhri.cn

* Correspondence: yanghanbo@tsinghua.edu.cn

Abstract: Raindrop size distributions (DSDs) are the microphysical characteristics of raindrop spectra. Rainfall characterization is important to: (1) provide information on extreme rate, thus, it has an impact on rainfall related hazard; (2) provide data for indirect observation, model and forecast; (3) calibrate and validate the parameters in radar reflectivity-rainfall intensity (Z-R) relationships (quantitative estimate precipitation, QPE) and the mechanism of precipitation erosivity. In this study, the one-year datasets of raindrop spectra were measured by an OTT Parsivel-2 Disdrometer placed in Yulin, Shaanxi Province, China. At the same time, four TE525MM Gauges were also used in the same location to check the disdrometer-measured rainfall data. The theoretical formula of raindrop kinetic energy-rainfall intensity (KE-R) relationships was derived based on the DSDs to characterize the impact of precipitation characteristics and environmental conditions on KE-R relationships in semi-arid areas. In addition, seasonal rainfall intensity curves observed by the disdrometer of the area with application to erosion were characterized and estimated. The results showed that after quality control (QC), the frequencies of raindrop spectra data in different seasons varied, and rainfalls with R within 0.5–5 mm/h accounted for the largest proportion of rainfalls in each season. The parameters in Z-R relationships ($Z = aR^b$) were different for rainfall events of different seasons (a varies from 78.3–119.0, and b from 1.8–2.1), and the calculated KE-R relationships satisfied the form of power function KE = AR^m, in which A and m are parameters derived from rainfall shape factor μ. The sensitivity analysis of parameter A with μ demonstrated the applicability of the KE-R formula to different precipitation processes in the Yulin area.

Keywords: raindrop size distribution; radar reflectivity; raindrop spectrometer; semi-arid area

---

1. Introduction

Characteristics of precipitation show the impact of meteorological conditions [1], and the measurement of quantitative distribution of precipitation is important for studying the mechanism of global climate and environmental change [2]. Raindrop size distributions (DSDs) show the microphysical properties of a rainfall event and vary with precipitation both in time and space. The DSDs are of importance to enhance the accuracy of quantitative precipitation estimation (QPE) by weather radar, and raindrop spectra have been used to calculate radar reflectivity factors in many studies [3–6]. Thus, investigating the raindrop spectra is essential for providing information on the

microphysical characteristics of different precipitation types, and improving the parameterizations of different rainfall processes.

Raindrop spectra can be measured by many methods, e.g., as done by Das et al., Waldvogel et al., Schönhuber et al. and Liu et al. [3,7–9], and these different methods vary in measurement principles and precision of data. In recent years, the raindrop disdrometer has been widely used to measure raindrop spectra because of its high measurement accuracy and small-time interval for data acquisition. Liu et al. [9] measured the precipitation in Nanjing, China and compared four different methods of rainfall to conclude that the disdrometer and other methods are consistent within the range of medium particle size. However, Zhang et al. [10] used three different methods to analyze the DSD characteristics in Zhuhai, China and found that the disdrometer had limitations in measuring small raindrops when compared to other methods. Raupach et al. [11] used a 2DVD device to correct the DSDs measured by three disdrometers and the correction showed its general applicability under different climate types. However, the study on the seasonal variation of rainfall characteristics in semi-arid areas of China using a raindrop disdrometer is very limited.

The measured raindrop spectra can be used to calibrate and validate the parameters in radar reflectivity-rainfall intensity (Z-R) relationships (quantitative estimate precipitation, QPE). Variability of DSDs in different forms of precipitation impact the radar reflectivity-rainfall intensity relationships (Z-R relationships, normally in the form of power function $Z = aR^b$, in which a and b are parameters derived from data fitting) [12,13], and the quantitative estimation of rainfall intensity (R) by Z-R relationships can be further modified. Sulochana et al. [14] investigated the Z-R relationships over a tropical station and concluded that the prefactor of Z-R relationships is larger for stratiform rain than for convective rain, which was in agreement with the results reported from two other tropical stations. Sulochana et al. also found that there were large variations in Z-R relationships in different seasons. Das et al. [3] analyzed the impact of different precipitation types on Z-R relationships in a hill station with a pronounced monsoon climate, and the results showed that the Z values of the shallow-convective system are the lowest, compared to other precipitation types. Das et al. concluded that the coefficient a is larger for stratiform rain than for convective and shallow-convective rain. However, there remains very limited research on using disdrometers to investigate raindrop spectra in semi-arid areas [15], and further research on the parameters in Z-R relationships is needed.

In addition, the measured raindrop spectra can be used to explore the mechanism of precipitation erosivity. The relationship between rainfall kinetic energy (KE) and intensity (R) is a significant approach to study the impact of precipitation on soil erosion [16], and a disdrometer can be deployed to measure the rainfall kinetic energy and intensity [17]. Angulo-Martínez et al. [18] measured and analyzed the uncertainty in KE-R relationships with five Parsivel disdrometers among three locations, and found that the types, accuracy and location of the disdrometers and precipitation types influence the estimation results of KE. Overestimation of the midsize raindrops led to a high estimation result of KE. Moreover, Carollo et al. [19] concluded that KE/R depends on the median volume diameter of precipitation events strictly, and this relationship does not rely on the locations of disdrometers. There have also been studies investigating the KE-R relationships in arid and semi-arid areas, e.g., Meshesha et al. and Abd Elbasit et al. [20,21], and many KE-R relationships were derived with this approach [22]. Nevertheless, many of the relationships vary greatly because of the different climate conditions, and a general formula for KE-R relationships needs to be derived to be suitably utilized in arid and semi-arid areas.

Yulin (in the northern region of Shaanxi Province, China) has a semi-arid climate, and the precipitation in this area is not evenly distributed throughout different seasons of the year [23]. The analysis of DSDs in different seasons is helpful to understand the variability of precipitation in this semi-arid area. The objectives of this paper are: (1) to analyze the detailed statistical data of DSDs based on the observation of a raindrop disdrometer located in Yulin, and collect data on the variability of microphysical characteristics of precipitation in different seasons; (2) to investigate the Z-R relationships in different seasons, and analyze the variability of the parameters in the Z-R relationships across

different precipitation types; (3) to derive a theoretical formula for KE-*R* relationships and further analyze the calculated results.

The following sections are organized below. Section 2 describes the research method, the datasets of the research area, and the derivation processes of theoretical formula. Section 3 analyzes the observed results statistically and presents the comparison of different precipitation periods in different seasons. Section 4 presents a further discussion of the comparison of disdrometer- and gauges-measured rainfall data, the comparison of different *Z-R* relationships and the theoretical formula of KE-*R* relationships through sensitivity analysis. Section 5 gives the conclusion.

2. Materials and Methods

2.1. Observation Station and Disdrometer Datasets

Microphysical characteristics of different precipitation types were measured at Yulin Ecohydrological Station (109°28′2.7′′E, 38°26′43.6′′N, 1236 m above sea level (a.s.l.), located in Mu Us Sandy Land) in Shaanxi Province, China. Mu Us Sandy Land has a semi-arid climate type with a low amount of precipitation [24]. The average annual precipitation in this area is 413 mm (from the year 1951 to 2018, measured by Yulin Meteorological Station (109°28′12′′E, 38°9′36′′N, 936 m a.s.l.)).

Figure 1a shows the location of Yulin Ecohydrological Station and Yulin Meteorological Station. One OTT Parsivel-2 Disdrometer was placed in Yulin Ecohydrological Station, and continuously recorded data during a 1-year period between 10 August 2018 and 10 August 2019. There are also four TE525MM rainfall gauges placed at the same location (shown in Figure 1b), and the specific placement of the disdrometer and the four gauges is proposed by Xie et al. [25], providing a reference for the measurement data by the disdrometer. Data were recorded every 30 min during the researching period, with the four gauges recording data simultaneously and individually. The disdrometer conducted a 54 cm^2 laser beam to record the rainfall spectrum every 1 min. The record of the disdrometer includes two parts: raindrop diameters (*D*) and raindrop terminal velocity (*v*). When precipitation particles pass the laser beam of the sensor, the beam is blocked off by the particles equal to the diameters, and the output voltage will be reduced. If there is no particle passing through, the voltage will then be recorded as maximum. The duration of the reducing signal will be used to determine the terminal speed of a particle. Through the observation by the disdrometer, radar reflectivity (*Z*), rainfall intensity (*R*) and kinetic energy (KE) are derived from *D* and *v*. Figure 2 summarizes how the materials and methods and results are managed via a schematic.

2.2. Processing Microphysical Datasets from Disdrometer

In order to process the calculation of the characteristic variables of precipitation, the formulas are derived as follows:

$$N(D_i) = \sum_{j=1}^{32} \frac{n_{ij}}{V_j \cdot \Delta D_i} = \sum_{j=1}^{32} \frac{n_{ij}}{A \cdot \Delta t \cdot v_j \cdot \Delta D_i} \tag{1}$$

where $N(D_i)$ (mm$^{-1} \cdot$ m^{-3}) is the number concentration of raindrops per unit diameter interval for raindrops per unit volume with the diameter equal to D_i (the *i*th-bin of diameters of the spectra); $A = 5.4 \times 10^{-4}$ m^2 is the sampling area scanned by the laser beam; $\Delta t = 60$ s is the sampling time interval; v_j is the raindrop terminal velocity of the *j*th-bin of velocities of the spectra; ΔD_i is the class spread of the *i*th-bin of diameters of the spectra.
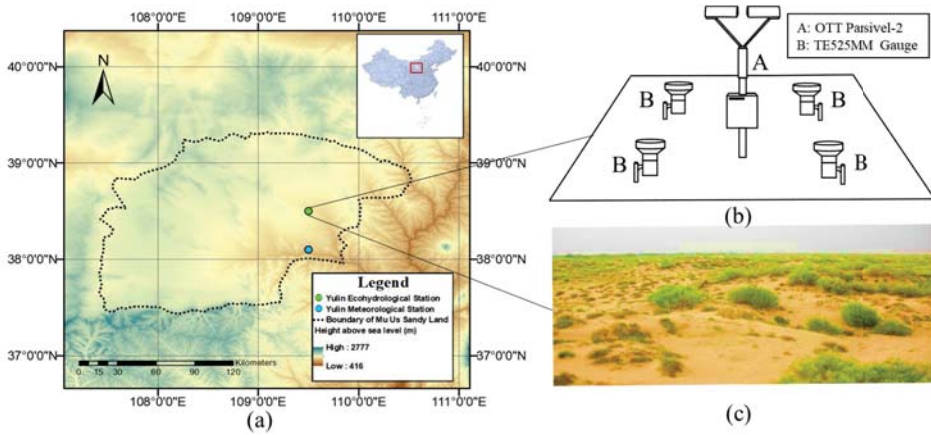
Figure 1. (**a**) The location of Yulin Ecohydrological Station (green point) and the location of Yulin Meteorological Station (blue point) and the DEM (Digital Elevation Model) characteristics around the sites. The red rectangle in the China map highlights the location of Mu Us Sandy Land; (**b**) the OTT Parsivel-2 disdrometer and the surrounding 4 TE525MM rain gauges in Yulin Ecohydrological Station; (**c**) vegetation cover situation in Yulin Ecohydrological Station.
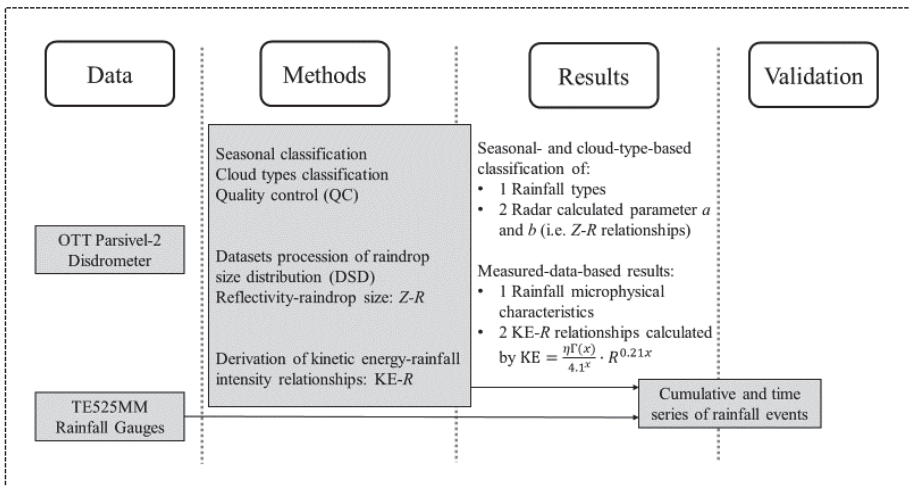
Figure 2. Schematic of the research technical route in this study.

The *n*th-moment of the DSD can be calculated as:

$$M_n = \int_0^\infty D^n N(D) dD \qquad (2)$$

Moreover the 6th-moment of the DSD is equal to the radar reflectivity Z ($\text{mm}^6 \cdot \text{m}^{-3}$):

$$Z = M_6 = \int_0^\infty D^6 N(D) dD = \sum_{i=3}^{32} D_i^6 N(D_i) \Delta D_i \qquad (3)$$

where the 1st-bin and the 2nd-bin of diameters are not evaluated in the measurements of the OTT Parsivel-2 because these two bins are out of the measurement range of the disdrometer, thus *i* starts from 3 to 32.

The $R_{\text{total}(i,j)}$ (mm·h^{-1}) is the total rainfall per unit time in the n_{ij} grid:

$$R_{\text{total}(i,j)} = v_j \cdot N(D_i) \cdot \Delta D_i \cdot \frac{1}{6}\pi D_i^3 \cdot 3600 \cdot 10^{-6} = \frac{3\pi}{5000}v_j D_i^3 N(D_i)\Delta D_i \tag{4}$$

The rainfall intensity R (mm·h^{-1}) is the summation of $R_{\text{total}(i,j)}$, and thus, can be calculated as:

$$R = \frac{3\pi}{5000}\sum_{i=3}^{32}(\sum_{j=1}^{32}v_j D_i^3 N(D_i)\Delta D_i) \tag{5}$$

The content of liquid raindrop W (g·m^{-3}) is the mass of liquid raindrops per unit volume:

$$W = \frac{\pi}{6}\cdot 10^{-3}\cdot \int_0^\infty D^3 N(D)dD = \frac{\pi}{6}\cdot 10^{-3}\cdot \sum_{i=3}^{32}D_i^3 N(D_i)\Delta D_i \tag{6}$$

The gamma distribution of raindrop spectra is derived as:

$$N(D) = N_0 D^\mu \exp{(-\Lambda D)} \tag{7}$$

where N_0 $(\text{mm}^{-1-\mu}\cdot\text{m}^{-3})$ is the intercept parameter and varies in dozens of orders of magnitudes [26]. Thus, the normalized intercept parameter Nw $(\text{mm}^{-1}\cdot\text{m}^{-3})$ is used by Testud et al. [27] to represent N_0, in order for the characteristics of the parameter of distribution of raindrop spectra can be calculated without any assumption about the DSD shapes:

$$N_w = \frac{4^4 M_3^5}{6\rho_w M_4^4} \tag{8}$$

N_w is better than N_0 for representing the raindrop concentration with certain raindrop sizes, as is not dependent on parameter μ. The mass-weighted mean diameter $Dm(\text{mm})$, and the rainfall shape parameter μ (dimensionless, related to rainfall types [28]), can be expressed as [29,30]:

$$Dm = \frac{4+\mu}{\Lambda} \tag{9}$$

$$\mu = \frac{1}{2\left(1 - \frac{M_4^3}{M_3^2 M_6}\right)}\left\{11\frac{M_4^3}{M_3^2 M_6} + \left[\frac{M_4^3}{M_3^2 M_6}\left(\frac{M_4^3}{M_3^2 M_6} + 8\right)\right]^{\frac{1}{2}} - 8\right\} \tag{10}$$

The slope parameter of gamma distribution Λ (mm^{-1}) is dependent on parameter μ, and Seela et al. [31] concluded that μ-Λ relationships vary with different precipitation types in different areas. In this study, the Λ is expressed as a function of the 2nd and 4th-moment of the DSD and parameter μ [32]:

$$\Lambda = \left[\frac{M_2\Gamma(\mu+5)}{M_4\Gamma(\mu+3)}\right]^{\frac{1}{2}} \tag{11}$$

2.3. Precipitation Types and Quality Control (QC)

In this study, in order to analyze the characteristics of raindrop spectra in different precipitation types, the precipitation data is classified into convective and stratiform rain based on the rainfall intensity [14]. Quality control is carried out in this study because there are errors in the measurement

of large diameters of raindrops using the disdrometer. In this study, raindrops larger than 6 mm in diameter were considered to have crushed during falling [33]. Therefore, raindrops larger than 6 mm in diameter are excluded in this study

Figure 3 illustrates the principle of classifying stratiform, convective, mixed-cloud rain and non-rainfall events. In this study, the effective observation range of raindrop diameter is 0.25–6 mm. Raindrop spectrum data with the total number of raindrops less than 10 or rain intensity less than 0.5 mm/h is regarded as noise [34,35].
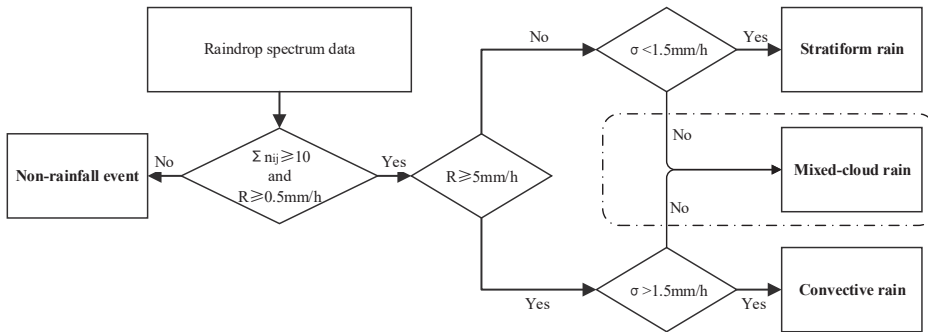
Figure 3. Classifying stratiform, convective rain and non-rainfall events. This classification method is derived by Li et al. [36] and used in Mt. Huangshan (118°10′E, 30°07′N, 1351 m a.s.l.), and is considered valid in this study to classify stratiform/convective rain.

The quality control principle is similar to the method used by Li et al. [36]. For an instantaneous moment t_n during a rainfall event, if the rain intensity R within the time range $[t_n - N_s, \ t_n + N_s]$ is > 5 mm/h and the standard deviation is >1.5 mm/h, the event is classified as a convective rainfall event; if the rain intensity R within the time range $[t_n - N_s, \ t_n + N_s]$ satisfies $0.5 < R < 5$ mm/h and the standard deviation is <1.5 mm/h, then the event is classified as a stratiform rainfall event. Other rainfall events are classified as mixed-cloud rain [37].

2.4. Derivation of KE-R Relationships

Rainfall kinetic energy (KE) can be calculated from the raindrop disdrometer, rainfall terminal velocity and raindrop size distribution. The total kinetic energy KE $(J/(m^3 \cdot s))$ can be derived as [25,38]:

$$KE = \frac{1}{12}\rho\pi a^3 N_0 \int_0^\infty D^{\mu+3+3b} e^{-\Lambda D} dD \qquad (12)$$

where $a = 3.78$ and $b = 0.67$ are constant parameters [39]. In arid or semi-arid areas, most of the precipitation is of weak or moderate levels and Marshall et al. [40] concluded that N_0 in the expression of $N(D)$ is almost fixed to 0.08 cm^{-4}. The parameter x is defined as:

$$x = \mu + 4 + 3b = \mu + 6.01 \qquad (13)$$

From Equations (12) and (13), the KE can be further expressed as:

$$KE = \frac{\rho\pi a^3 N_0}{12\Lambda^x} \int_0^\infty t^{x-1} e^{-t} dt = \frac{\rho\pi a^3 N_0 \Gamma(x)}{12\Lambda^x} \qquad (14)$$

The expression of Λ is:

$$\Lambda = 4.1 R^{-0.21} \qquad (15)$$

which is commonly used in the calculation of KE-*R* relationships in many studies [41–43], and:

$$KE = \frac{\rho \pi a^3 N_0 \Gamma(x)}{12 \cdot 4.1^x} \cdot R^{0.21x} \tag{16}$$

The parameter η is defined as:

$$\eta = \frac{\rho \pi a^3 N_0}{12} \tag{17}$$

and is constant. Then:

$$KE = \frac{\eta \Gamma(x)}{4.1^x} \cdot R^{0.21x} \tag{18}$$

Therefore, the empirical formula of the relationship between KE and *R* with the variable parameter *x* for the semi-arid area is derived.

3. Results

3.1. Data after Quality Control

In this study, 8184 records are selected as rainfall events, including 7388 stratiform and 796 convective data. The amount of observed mixed-cloud rain is zero. Figure 4 shows the frequency accumulation curve of rainfall intensity recorded. The average $\overline{R}$ of the recorded data is 2.3 mm/h (the calculated average $\overline{R}$ includes only rainy hours, as shown in Section 2.3), and the data with rainfall intensity less than 5 mm/h is more than 90% of the total time of rainfall data.
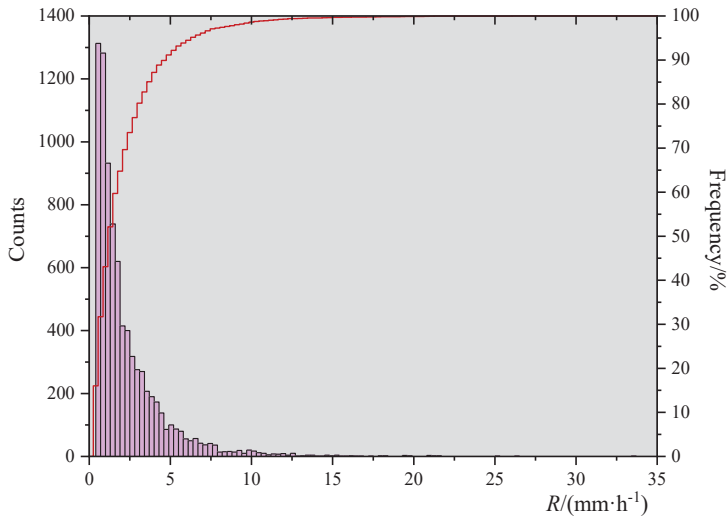
Figure 4. Frequency distribution of rain rates calculated from the whole OTT Parsivel-2 disdrometer datasets.

A division of the rainfall data into different seasons (spring from March to May, Summer from June to August, Autumn from September to November and Winter from December to February) and rainfall types (stratiform and convective rain) acquired by the OTT Parsivel-2 is reported in Table 1. The results shown in Table 1 reflect the trend of precipitation change in this region with different seasons. The frequency of raindrop spectra data recorded in summer is the highest (with 7019 records), followed by autumn (4230 records), spring (3387 records) and winter (only 116 records). In spring, summer and autumn, the total time of stratiform rainfall events T_s is higher than the total time of convective rainfall events T_c, which reflects the precipitation characteristics of semi-arid areas [44].

The percentage of stratiform rain is from small to large summer (85.9%), spring (93.5%) and autumn (95.5%). Winter is not considered when comparing the characteristics of different seasons due to too few liquid precipitation data recorded. The percentage of convective rain, ordered from small to large, is spring (4.5%), autumn (6.5%) and summer (14.1%). In summer the R varies most and has the highest standard deviation of 2.8 mm/h. However, the percentage of rainfall time over the seasons is ordered from small to large, are spring (1.4%), autumn (1.7%) and summer (3.0%).

Table 1. Disdrometer-measured data of different seasons: total time with recorded data, the percentage of stratiform and convective rain and percentage of different ranges of rainfall intensity; maximum and average rainfall intensity, and standard deviation of rainfall intensity.

| Rainfall Characteristics	Spring (March–May)	Summer (June–August)	Autumn (September–November)	Winter (December–February)
Total time with data recorded T_{total} (min)	3387	7019	4230	116
Total time of stratiform rain T_s (min)	1813	3431	2126	18
Total time of convective rain T_c (min)	86	563	147	-
Percentage of stratiform rain $T_s/(T_s+T_C)$ (%)	95.5	85.9	93.5	100
Percentage of convective rain $T_c/(T_s+T_c)$ (%)	4.5	14.1	6.5	-
Percentage of rainfall time over the season (%)	1.4	3.0	1.7	≈ 0
Maximum rainfall intensity R_{max} (mm·h^{-1})	12.6	33.7	21.2	1.1
Average $\overline{R}$ (mm·h^{-1})	2.0	2.6	2.0	0.7
Median R_m (mm·h^{-1})	1.9	1.9	1.6	0.7
Precipitation Accumulation P_{total} (mm)	76.7	236.9	105.9	- *
Standard deviation of rainfall intensity R (mm·h^{-1})	1.6	2.8	1.8	0.15
$T_{0.5-2\ mm·h^{-1}}/(T_s+T_C)$ (%)	60.3	58.8	66.4	100
$T_{2-5\ mm·h^{-1}}/(T_s+T_C)$ (%)	35.1	27.1	27.1	-
$T_{5-10\ mm·h^{-1}}/(T_s+T_C)$ (%)	4.2	11.3	6.0	-
$T_{>10\ mm·h^{-1}}/(T_s+T_C)$ (%)	0.4	2.8	0.5	-

* The rainfall accumulation is not calculated in winter because of the inclusion of solid precipitation (e.g., snow).

Weak and moderate rainfalls (with rainfall intensity satisfying $0.5 < R < 5$ mm/h) account for the largest proportion of rainfalls in each season, and data satisfying $R > 5$ mm/h in summer is of the highest percentage of total rainfall time in the season (11.3%). The R_{max} is 33.7 mm/h in summer, 21.2 mm/h in autumn and 12.6 mm/h in spring. The $\overline{R}$ and P_{total} in summer are also the highest (2.6 mm/h and 236.9 mm, respectively), while the difference in $\overline{R}$ between spring (2.0 mm/h) and autumn (2.0 mm/h) is not significant. The difference in R_m between spring (1.9 mm/h) and summer (1.9 mm/h) is not significant.

3.2. Microphysical Characteristics of Precipitation in Different Seasons

Figure 5 is the distribution of different diameters with varying velocities obtained in different seasons. In spring, summer and autumn, the majority of raindrop particles are in an area close to the theoretical curve proposed by Beard [45]; in winter, the data is accumulated in low levels of both v and D. Data are overall lying over the theoretical curve in spring, summer and autumn; in winter, however, the velocities are underestimated compared to the theoretical curve in winter. This is because the type of rainfalls during winter was mainly snow (regarded as solid rainfalls) and could cause deviation in call speed, and the raindrop data with D > 6 mm are excluded as shown in Section 2.3.
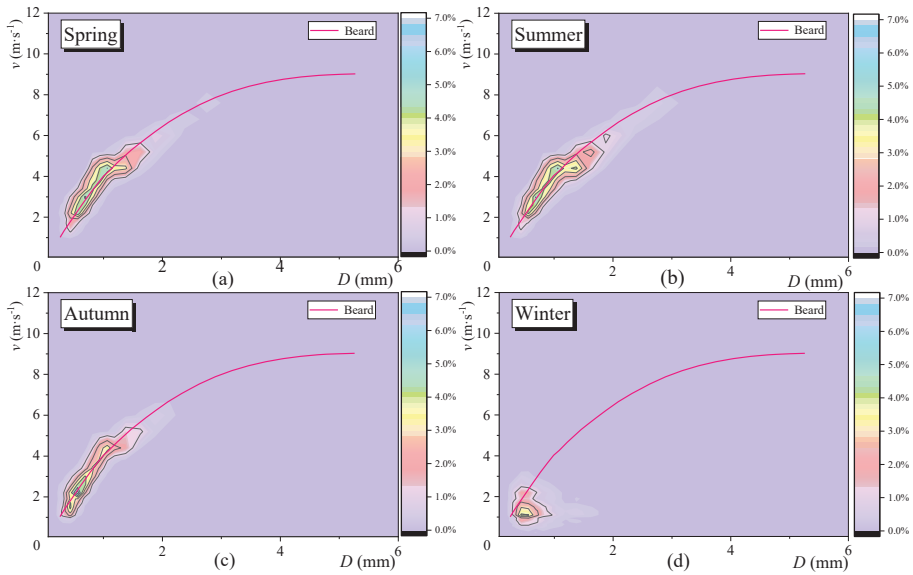
Figure 5. Plots of measured raindrop terminal velocity-diameter relationships in different seasons. The red curve is derived by Beard [45] in 1976, giving the relationships between *v* and *D*, and used as the reference line for the *v-D* distribution in each season (**a**) spring; (**b**) summer; (**c**) autumn; (**d**) winter.

Figure 6 shows the histogram of *Dm* and $\log_{10}Nw$ in different seasons and precipitation types. The yearly average *Dm* and $\log_{10}Nw$ are 1.41 and 3.91 mm, respectively. The average *Dm* of stratiform (convective) rain from small to large is 1.38 mm (1.46 mm) in summer, 1.46 mm (1.52 mm) in spring and 1.51 mm (1.57 mm) in autumn. However, there is less variation in average $\log_{10}Nw$ data among the three seasons. The average $\log_{10}Nw$ of stratiform rain from small to large is 3.88 mm in summer, 3.89 mm in autumn and 3.92 in spring. However, the maximum of the average $\log_{10}Nw$ of convective rain is 4.28 in summer. This reflects the micro-physical characteristics of rainfall in Yulin area: according to the results shown in Section 3.1, the average rainfall intensity is larger in summer than in spring and autumn. According to Equation (5), the rain intensity is related to raindrop diameter and DSD. The average *Dm* of different precipitation types in different seasons shows that for stratiform rain, *R* is affected more by the raindrop diameter; for convective rain, *R* is affected more by DSD. The average value of *Dm* is slightly less than that of in southern China (1.46 mm) [10].

Besides the average value, standard deviation and skewness of different *Dm* and $\log_{10}Nw$ were also calculated. The standard deviation of *Dm* in stratiform and convective rain among different seasons varies from 0.19 mm to 0.26 mm. The skewness of *Dm* and *Nw* in stratiform and convective rain are less than 0 in spring and autumn, illustrating the frequency of the data below the *Dm* mean (*Nw* mean) is less than data above the *Dm* mean (*Nw* mean). However, the skewness of *Dm* and *Nw* in stratiform and convective rain in summer are larger than 0, illustrating the frequency of the data below the *Dm* mean (*Nw* mean) is more than data above the *Dm* mean (*Nw* mean).
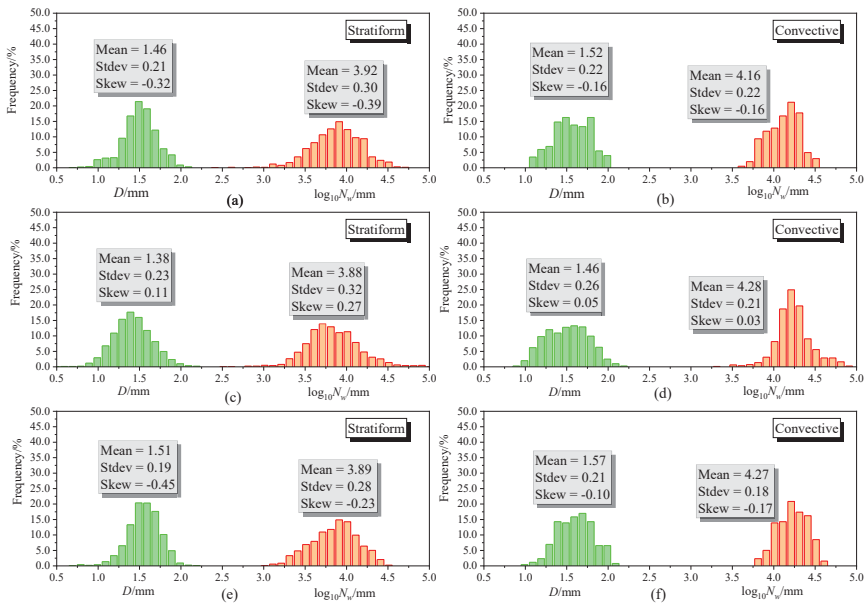
Figure 6. Frequency histogram of mass-weighted median diameter *Dm* and the denary logarithm of *Nw* in different precipitation types and different seasons calculated from the data measured by OTT Parsivel-2. (**a**) Histogram in stratiform rain in spring; (**b**) Histogram in convective rain in spring; (**c**) Histogram in stratiform rain in summer; (**d**) Histogram in convective rain in summer; (**e**) Histogram in stratiform rain in autumn; (**f**) Histogram in convective rain in autumn. The average value, standard deviation and skewness are also given for *Dm* and *Nw* in each plot.

Figure 7 shows the *Nw-R* relationships in different seasons and precipitation types. The base-10 logarithm of *Nw* is used to fit the relationship curves $\log_{10}Nw = cR^d$, in which c and d are parameters fitted by measured data. For the error bars in each panel, *R* in the range (0.5, 5) (for stratiform rain) are divided into nine intervals evenly, and in the range (5, 35) (for convective rain) are divided into five intervals ($5 < R < 10$, $10 < R < 15$, $15 < R < 20$, $20 < R < 25$ and $R > 25$ mm·h^{-1}), and error bars are used for each interval. The error bars for each interval are based on the mean value of *R* and log10*Nw*, with the ±1 Stdev (standard deviation), respectively. A significance analysis of fitting results is also proposed [46]. The p-values in each panel of Figure 7 are derived from the fitting tests of the power function. The p-values show that the fits for stratiform rain are statistically relevant and sound (shown in Figure 7a,c,e). When comparing the disparity in precipitation types, parameters c and d have a smaller range in variability in different seasons. Figure 7a,c,e indicate that for stratiform rain, the difference of parameter c among different seasons ranges in 3.73–3.79 ($p < 0.05$). However, the parameter c for convective rain varies in a larger range of 3.86–4.14 ($p < 0.1$). For each season, the parameter a of stratiform rain is smaller than that of convective rain. The difference in parameter d is small among different seasons, and the d of stratiform rain is larger than that of convective rain.
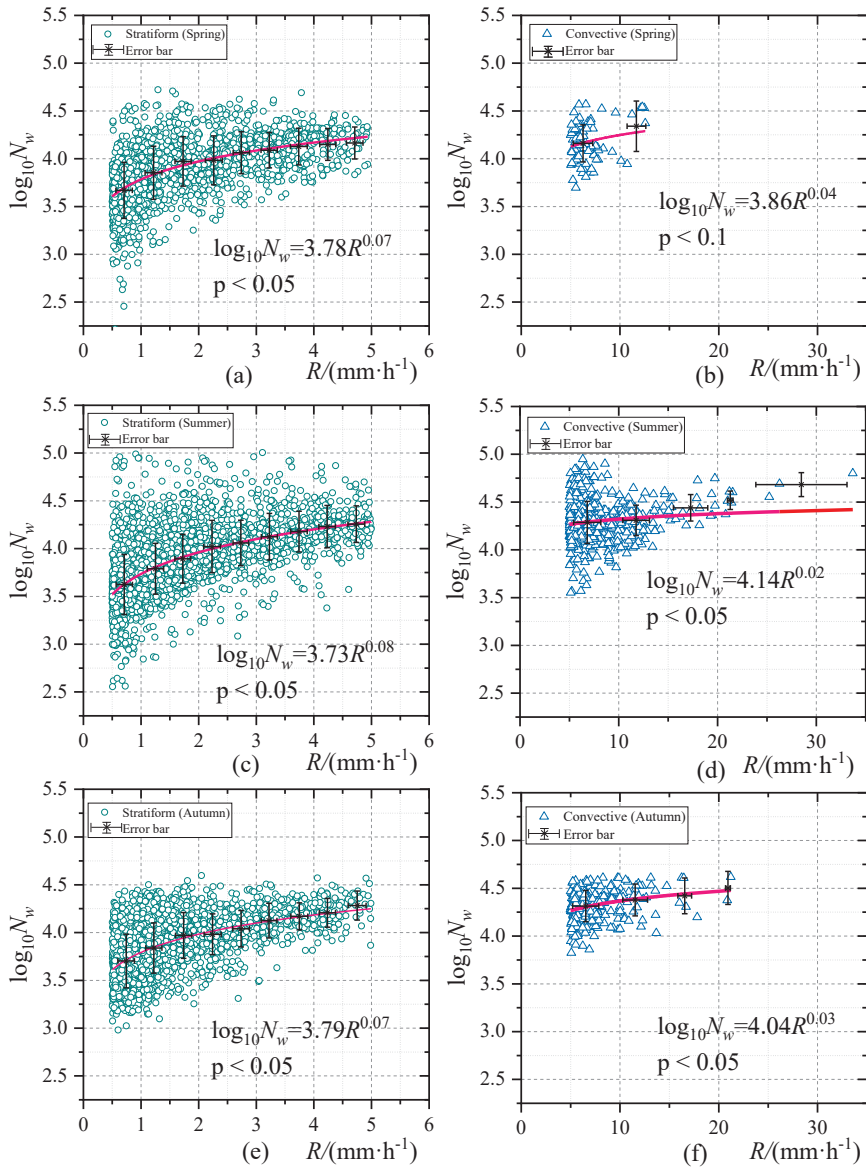
Figure 7. *Nw-R* relationships for different precipitation types in different seasons. The fitted power formula based on the least-squares method is also shown in each plot: (**a**) stratiform rain in spring; (**b**) convective rain in spring; (**c**) stratiform rain in summer; (**d**) convective rain in summer; (**e**) stratiform rain in autumn (**f**) convective rain in autumn.

3.3. Z-R Relationships in Different Rainfall Events

Three rainfall events were selected to illustrate the differences between the data measured by OTT Parsivel-2 Disdrometer and four TE525MM Rainfall Gauges. The three events are, respectively, chosen from each season, regarding their representation of different levels of rainfall intensity (all of the three events last for > 4h and include stratiform and convective records).

Table 2 shows the precipitation characteristics of these rainfall events. S/C records show the relative rates of amounts of the stratiform and the convective records during the selected rainfall events, for which Event 1, 2 and 3 are 662%, 888% and 753%, respectively. The results correspond to the average rainfall intensity $\overline{R}$ of each rainfall event. Event 2 is with the least $\overline{R}$ (2.42 mm·h^{-1}), while Event 1 is with the highest $\overline{R}$ (3.07 mm·h^{-1}).

Table 2. Precipitation characteristics of selected three rainfall events.

Rainfall Number	Rainfall Time	Season	$\overline{R_d}$/(mm·h^{-1}) *	$\overline{R_g}$/(mm·h^{-1}) **	Duration	S/C Records
Event 1	19–20 April 2019	Spring	3.36	3.51	4 h 30 min	225/34
Event 2	3 August 2019	Summer	1.65	1.78	12 h 30 min	897/101
Event 3	1 September 2018	Autumn	3.09	3.22	4 h	241/32

* Derived from disdrometer-measured raindrop spectra. See in Equation (5). ** Obtained by the four rainfall gauges and averaged.

Figure 8 shows the difference in *Z-R* relationships of the three rainfall events of spring, summer and autumn, and the results are $Z = 109.6R^{2.1}$ (Event 1), $Z = 119.0R^{1.8}$ (Event 2) and $Z = 78.3R^{1.9}$ (Event 3). The parameter a of each of the three rainfall events is less than the parameter a of the default *Z-R* relationship [47] ($Z = 300R^{1.4}$, where parameter a is equal to 300). Parameter b in $Z = aR^b$ has only slight difference among different rainfall events, and ranges in 1.8–2.1 ($p < 0.05$). This is consistent with previous studies [48–51], that parameter b in $Z = aR^b$ varies in the range of 1–2.87 ($p < 0.05$). The raindrop spectra have obvious changes for different precipitation types, causing the parameters of *Z-R* relationships to vary. It is clear that the *Z-R* relationships of rainfall events vary depending on the season and the precipitation types. The rainfall events with a higher S/C rate tend to have a higher average rainfall intensity. Therefore, precipitation estimates for different types should be treated as such. The results indicate that there is a need to utilize modified *Z-R* relationships in different seasons when calculating the rainfall intensity by the QPE method.

3.4. KE-R Relationships

The rainfall kinetic energy KE and rainfall intensity *R* are calculated based on Equations (5) and (18), respectively. The KE-*R* relationships of stratiform rain in different seasons are derived in the form of a power function. The fitting results show the KE-*R* relationships satisfy the form of power function $KE = AR^m$, in which *A* and *m* are parameters.

Figure 9 shows that the stronger the rainfall intensity is, the faster the rainfall kinetic energy tends to increase. The exponent of different power functions m varies from 1.45 to 1.82, in order from small to large, spring < autumn < summer. Parameter *A* has similar values in spring (equal to 8.56) and in summer (equal to 8.10), but in autumn the parameter *A* is 5.56. The value of the parameter *A* is closely related to the raindrop spectra [52]. The parameter *A* is also related to the median rainfall intensity R_m, which results in the decrease of *A* in autumn compared to spring and summer.
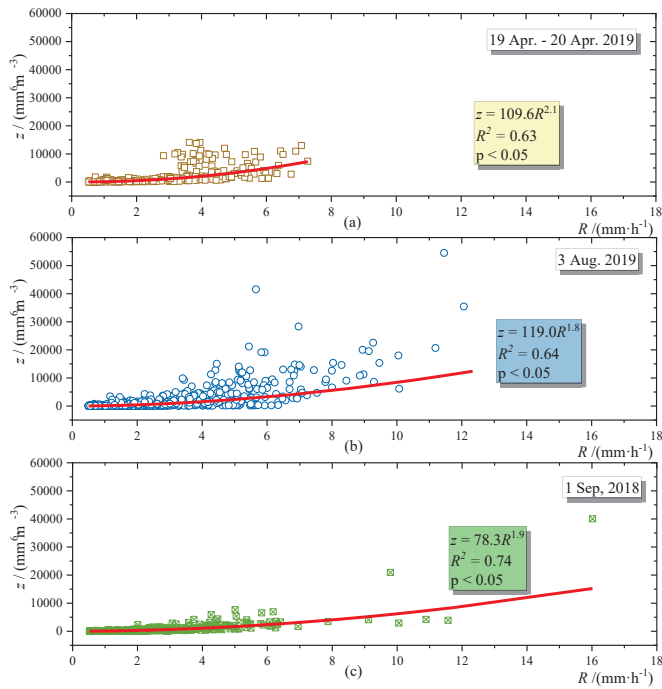
Figure 8. *Z-R* relationships ($Z = aR^b$) in three rainfall events. Each rainfall event is selected from different seasons (spring, summer and autumn). (**a**) *Z-R* relationship of a rainfall event in spring, from 19 April to 20 April 2019; (**b**) *Z-R* relationship of a rainfall event in summer, on 3 August 2019; (**c**) *Z-R* relationship of a rainfall event in autumn, on 1 September 2018; the fitted curves are derived based on the power function, which are solid red curves in (**a–c**).
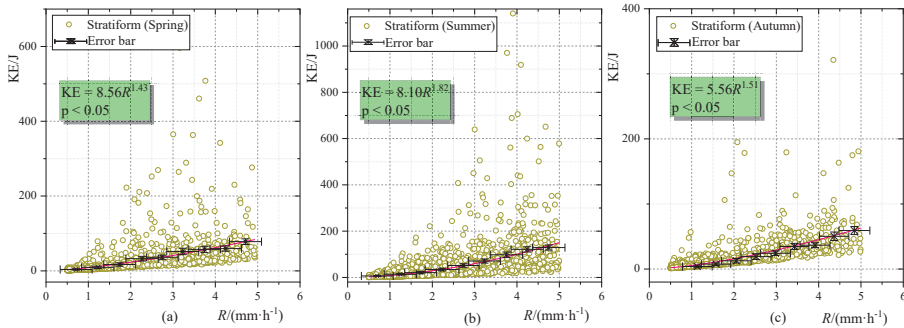
Figure 9. KE-*R* curves fitted in different seasons. Data of stratiform rain in spring, summer, and autumn are analyzed. (**a**) KE-*R* relationship in spring; (**b**) KE-*R* relationship in summer; (**c**) KE-*R* relationship in autumn. For the error bars in each panel, *R* in the range (0.5, 5) are divided into nine intervals evenly, and error bars are used for each interval. The error bars for each interval are based on the mean value of *R* and KE, with the ±1 Stdev (standard deviation), respectively.

4. Discussion

4.1. Uncertainty Analysis of Disdrometer-Measured Data

Figure 10 shows the relationship between two different measurements on a daily scale. Each point represents the total rainfall of a given day during the 1-year observation period. The abscissa is the total measured rainfall of Parsivel-2, and the ordinate is the averaged total rainfall measured by four TE525MM Gauges (data with both daily total rainfall greater than 1mm are taken for analysis). The results show that the disdrometer-measured data demonstrates a good correlation with the average rainfall gauge-measured data, and the deviation is mainly caused by the instrument theoretical error in measurement of disdrometer and gauges (e.g., sampling error), which was concluded by Tokay et al. [29]. The measured rainfall data of the disdrometer is suitable for this semi-arid area.
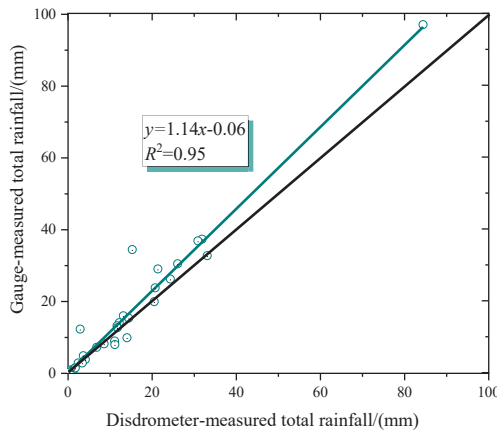
Figure 10. Scatter plot of gauge- and disdrometer-measured total rainfall on a daily scale. The results obtained by the two measurement methods are compared with the scatter plot. The solid green line is the fitted result.

Figure 11 shows the precipitation process of three rainfall events, with the calculated average of the data measured by four rainfall gauges. The determinate coefficient R^2 of a linear relationship between Parsivel Disdrometer and TE525MM Gauges is 0.68 for (a), 0.83 for (b) and 0.96 for (c). The standard deviations of disdrometer-measured (gauge-measured) rainfall intensity in (a), (b) and (c) are 1.37 mm/h (2.90 mm/h), 1.02 mm/h (1.43 mm/h) and 1.43 mm/h (2.46 mm/h), respectively. The standard deviation of the disdrometer-measured data in each season is less than the gauge-measured data in the corresponding season. In order to further compare the difference between disdrometer-measured and gauges-measured data, the mean absolute error(MAE) and the root mean square error(RMSE) of the events are also given. Event 3 has the smallest MAE (0.69 mm·h^{-1}) and smallest RMSE (0.99 mm·h^{-1}) among the three events. This indicates that the disdrometer can potentially estimate the precipitation with better performance during Event 3, according to the comparison with gauge-measured data.
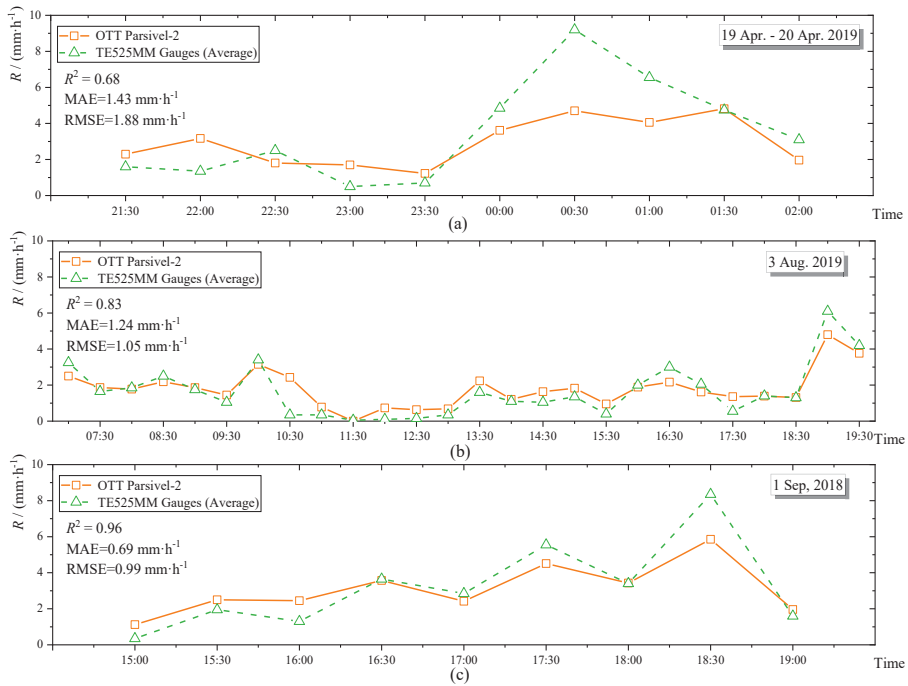
Figure 11. The comparison of measurement results of OTT Parsivel-2 (orange solid lines) and four TE525MM gauges (green dotted lines, averaged) in 3 rainfall events. (**a**) a rainfall event in spring, from 19–20 April 2019; (**b**) a rainfall event in summer, on 3 August 2019; (**c**) a rainfall event in autumn, on 1 September 2018.

The measuring range of Parsivel-2 is 0.25–6 mm, and it has been proved that the disdrometer has errors in measuring both smaller and larger diameters of raindrops in many studies [53,54], which have more accuracy in measuring mid-level raindrops in the spectra. Liu et al. [54] further explained the algorithm of the disdrometer and concluded that the differences can be attributed to a certain error of rainfall amount accumulation. In all three measured events, when the TE525MM Gauges-measured data reaches its peak (at 0:30 in (a), 19:00 in (b), and 18:30 in (c)), the data measured by gauges are all more than the disdrometer-measured data; and when the TE525MM Gauges-measured data reaches its minimum of the event (at 23:30 in (a), at 11:30 in (b) and 15:00 in (c)), the data measured by gauges are all no more than the disdrometer-measured data. As raindrop spectra data larger than 6 mm are treated as solid rainfall and eliminated, the peak value of the disdrometer is lessened. This is consistent with the results in Section 3.3 that in Event 1 the disdrometer- and gauge-measured $\overline{R}$ data indicates a significant gap (2.94 and 3.51 mm·h^{-1}, respectively). Since the measured raindrop spectra data larger than 6 mm have been removed, and raindrops of large diameter are often accompanied by heavy rain [25], it is speculated that this is the reason why the peak of rainfall intensity obtained by the raindrop disdrometer is smaller than the value measured by rain gauges.

4.2. Comparison of Different Z-R Relationships

In this section, the *Z-R* relationships of different experiment areas are compared, as shown in Table 3. The *Z-R* relationships calculated in different studies vary with different measurement methods, different study areas and rainfall time and types. Previous studies have shown that parameter a has a very wide range of values. This is consistent with Kang [55] that the range of parameter a is 10–1200. The parameters a and b vary with different seasons and locations, as well as with the measurement methods. It can be concluded that there is a great variation in the parameters a and b at different places, and *Z-R* relationships should be determined according to different regions. The above analysis also shows that it is of great significance to further improve the accuracy of radar rain measurement inversion and to further study the microphysical characteristics of rainfall [56].

Table 3. Results of *Z-R* relationships in different study areas.

Reference [57–62]	Precipitation Time/Season	Location	Measurement Method	Z-R relationships	Rainfall Type
Zhang et al., 1992	July in 1981	Jiangsu, Eastern China	Digital Radar	$Z = 60R^{2.3}$	Heavy rain
Marshall et al., 1947	Summer in 1946	Canada	MHF Radar	$Z = 190R^{1.72}$	Thunderstorm
Chen et al., 2008	15 July–7 August in 2007	Leizhou Peninsula, China	Parsivel Disdrometer	$Z = 269R^{1.54}$ $Z = 178R^{1.58}$	Convective rain Mixed-cloud rain
Marzuki et al., 2013	2006–2007	Kototabang, West Sumatra, Indonesia	2D-Video Disdrometer	$Z = 136R^{1.26}$	Stratiform rain
Hazenberg et al., 2011	14 September 2006	Cévennes-Vivarais region, France	OTT/Parsivel Disdrometer	$Z = 79R^{1.52}$	Stratiform rain
Blanchard, 1953	October in 1951– August in 1952	Hawaii, USA	Filter-paper	$Z = 31R^{1.71}$	Thunderstorm
Sivaramakrishnan, 1961	November in 1958	India	Filter-paper	$Z = 67.6R^{1.94}$	Thunderstorm

4.3. Sensitive Analysis of the Formula of KE-R Relationships

Many previous studies [22,52,63] have shown that parameters of the KE-*R* relationship are highly sensitive to DSD, and for parameters closely related to DSD characteristics, such as μ, the change in μ represents the difference in rainfall characteristics, which can affect the value of the parameters. Equation (18) indicates the KE-*R* relationships under the impact of parameter *x*. The parameter *A* and m are used to simplify the KE-*R* formula (expressed as Equation (18)) and are defined as:

$$A = \frac{\eta \Gamma(x)}{4.1^x} \tag{19}$$

$$\text{m} = 0.21x \tag{20}$$

in which *A* and m are functions of parameter *x*, thus, *A* and m are affected by raindrop microphysical characteristics and the environmental conditions. The parameter *x* (known as a linear variation with the parameter μ according to the Equation (13)) varies differently due to the different regions, different time periods. Therefore, sensitivity analysis is conducted on the parameters *A* and m with the change of parameter *x*, in order to indicate the applicability of the KE-*R* relationships formula (derived as Equation (18)) under different rainfall conditions. S_1 and S_2 were defined as the parameters to analyze the sensitivity of *A* and m with the change of *x*, and the influence of change on parameter *x* on the KE-*R* relationships is further shown. The definition of sensitive parameters S_1 and S_2 are derived in Equations (21) and (22):

$$S_1 = \lim_{\Delta x \to 0} \left(\frac{\Delta A/A}{\Delta x/x} \right) = \frac{dA}{dx} \cdot \frac{x}{A} \tag{21}$$

$$S_2 = \lim_{\Delta x \to 0} \left(\frac{\Delta \text{m}/\text{m}}{\Delta x/x} \right) = \frac{d\text{m}}{dx} \cdot \frac{x}{\text{m}} = 1 \tag{22}$$

in which S_1 (S_2) is the ratio of the change rate of parameter A (parameter m) to the change rate of parameter x. It is obvious from Equation (20) that the parameter m is proportional to parameter x, thus, S_2 is a constant, as shown in Equation (22). The sensitive parameter S_1 is further calculated of stratiform and convective rain obtained via disdrometer in different seasons, and the results are summarized in Table 4.

Table 4. Sensitive analysis results for parameter A with x of different precipitation types in different seasons and the total year.

	Spring	Summer	Autumn	Total Year
S_1 (Stratiform)	4.10	4.68	3.83	4.29
S_1 (Convective)	3.88	3.79	3.97	3.88

Table 4 shows the results of the sensitivity analysis for A with x of different precipitation types in different seasons. The sensitivity of parameter A varies with different precipitation types and seasons. In the whole-year scale, the S_1 is 4.29 of stratiform rain and 3.88 of convective rain, which indicates the parameter A of stratiform rain is more sensitive to the change of parameter x in the total year. The S_1 of stratiform rain in spring and summer is also more than that of convective rain in corresponding seasons. However, the S_1 of stratiform rain (3.83) is less than that of convective rain (3.97) in autumn, and the S_1 reaches its lowest in autumn among three seasons. The parameter x is affected by changes in environmental factors and rainfall types [64], and in autumn, the sensitivity of parameter A with the change of x is lower than in spring or summer. This explains the obvious decrease of parameter A in autumn compared with the other two seasons calculated in Section 3.4. For convective rain in different seasons, the S_1 from small to large is summer (3.79) < spring (3.88) < autumn (3.97).

According to Teng [65], the values of μ for raindrop spectra are between -1 and 4, and different μ values correspond to different rainfall characteristics. The larger the μ, the more likely it is to cause convective rain. In order to analyze the KE-R relationship when the rainfall shape parameter μ takes different values, the KE-R relationship curve of different values of μ was therefore made based on Equation (18), as shown in Figure 12a–f. The rule of this study is that both parameter A and m increase with the increase of μ. Figure 12g further shows the relationship among KE-R-μ. At a certain R, KE increases as μ increases. At the same rainfall intensity, the rainfall kinetic energy is also related to the rainfall type: the more the rainfall type is inclined to convective rain, the greater the rainfall kinetic energy will be, which can be explained by the Equation (18) and corresponds to the conclusions in other studies [66,67]. KE-R relationship changes with the change of the parameter x, according to Equation (18). From Equation (13), the parameter x can be calculated with a linear function from parameter μ. As discussed in Section 2.2, the parameter μ is correlated with the rainfall types; thus, the KE can be interpreted by the rainfall types according to Equation (18). In addition, different fitted formulas obtained in previous studies can be approximated with Equation (18) by changing the parameter μ (see Table A1), and the specific results can be approximated by Figure A1. This shows that the derived theoretical formula (Equation (18)) in this study is universal in various regions; however, whether the formula can be directly used to further analyze the KE-R relationship in other semi-arid areas should be further discussed in the future.
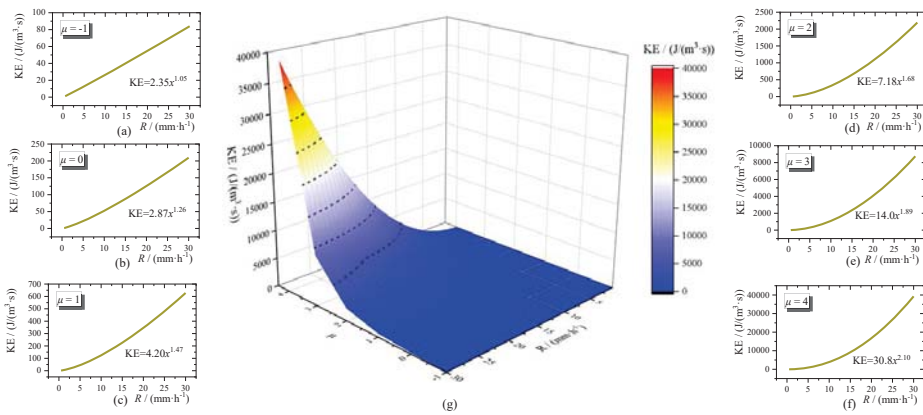
Figure 12. The *R*-KE relationships under different values of shape factors *μ*. The curves are derived based on Equation (18). (**a**) KE-*R* relationship for *μ* = −1; (**b**) KE-*R* relationship for *μ* = 0; (**c**) KE-*R* relationship for *μ* = 1; (**d**) KE-*R* relationship for *μ* = 2; (**e**) KE-*R* relationship for *μ* = 3; (**f**) KE-*R* relationship for *μ* = 4; (**g**) KE-*R*-*μ* relationship.

5. Conclusions

Characteristics of raindrop size distributions (DSDs) are important for improving the accuracy of radar reflectivity-rainfall intensity (*Z*-*R*) relationships in remote sensing (QPE) and the estimation of soil erosivity. In this study, an OTT Parsivel-2 Disdrometer is used to measured raindrop spectra from 10 August 2018 to 10 August 2019 in Yulin Ecohydrological Station, Shaanxi Province, China. The precipitation events obtained are classified as stratiform and convective rain based on the rainfall intensity classifying processes. The conclusions are summarized as follows.

(1) The characteristics of microphysical variables (the mass median diameter *Dm* and the raindrop size distribution *Nw*) were analyzed. The average *Dm* of different precipitation types in different seasons shows that for stratiform rain, rainfall intensity *R* is affected more by the average raindrop diameter *Dm*; for convective rain, *R* is affected more by DSD. The yearly average *Dm* and $\log_{10}Nw$ are 1.41 and 3.91 mm, respectively. The average *Dm* of stratiform (convective) rain from small to large is 1.38 mm (1.46 mm) in summer, 1.46 mm (1.52 mm) in spring and 1.51 mm (1.57 mm) in autumn. This reflects the semi-arid climate rainfall characteristics in Yulin Station.

(2) The variances of rainfall microphysical characteristics in different precipitation types and seasons are related. The distribution of rainfall terminal velocity-diameter (*v-D*) spectra of spring, summer and autumn is concentrated near the theoretical curve derived by Beard [45]. The base-10 logarithm of *Nw* is used to fit the relationship curves $\log_{10}Nw = cR^d$, in which c and d are parameters fitted by measured data. The difference in parameter d is small among different seasons (0.07–0.08 for stratiform rain and 0.02–0.04 for convective rain), and the d of stratiform rain is larger than that of convective rain.

(3) The *Z*-*R* relationships of different rainfall events in spring, summer and autumn in this semi-arid area are derived in this study. The parameter a is larger in stratiform rain than in convective rain, while the parameter b is larger in convective rain, showing the impact of different rainfall types on a and b. The results show that the estimation of different seasons should be treated, respectively.

(4) The theoretical formula of KE-*R* relationships for stratiform precipitation in semi-arid areas is derived (KE $= \frac{\eta\Gamma(x)}{4.1^x}\cdot R^{0.21x}$, where the parameter *η* is constant and $x = \mu + 6.01$), which indicates the characteristics of precipitation and environmental conditions represented by parameter *μ*. This formula gives a general expression of the KE-*R* relationships and is simple to use because the parameters are all derived from the parameter *μ*. The sensitivity analysis results show that the parameter *A* for stratiform rain is more sensitive to the change of different precipitation types and environmental conditions in a

total year. The closer the precipitation types are to convective rain, the larger the KE is at the same level of *R*. By changing the parameter *μ*, different empirical formulas obtained in previous studies can be approximated with the derived theoretical formula (Equation (18)).

In summary, the DSD characteristics of Yulin Station were obtained and the results can help to understand the microphysical characteristics of precipitation and have a strong impact on the mechanism of soil erosivity in the semi-arid area. Additionally, the formula of KE-*R* relationships provides a convenient way to fit with different rainfall events in semi-arid areas by adjusting its parameters. But the results are not conclusive because of the limited sample records of different rainfall types. In this study, data in winter is not deeply investigated, e.g., the solid precipitation processes should be further considered and analyzed. Moreover, the impact of environmental conditions on the parameter *A* and m is still not well understood. In the future, the influence of environmental factors on the parameters in *Z-R* relationships should be further discussed.

Appendix A

Table A1 shows different fitting formulas of KE-*R* relationships in different research areas. KE-*R* formulas obtained in different regions vary, including polynomial, exponential, logarithm and other forms. In order to reflect the universality of the theoretical formulas introduced in this paper, KE-*R* relationships are obtained by changing equations based on Equation (18) to approximate the formulas given in different studies. The deterministic coefficient is used as the index to evaluate the approximation results, and the fitting results ensure that the deterministic coefficient of the two is greater than 0.99 for each reference listed in Table A1. The scatter points obtained by the KE-*R* relationship obtained by Equation (18) are described, and then, the linear fitting is performed with the points traced based on the original equation; results are shown in Figure A1. The abscissa is the KE obtained by changing different equations in this paper, and the ordinate is the relation obtained by different researchers. The red line is the fitting result. The rainfall intensity corresponding to the horizontal and vertical coordinates of each point on the line is the same. Figure A1 shows that the deterministic coefficients of different results are all above 0.99, indicating that Formula (18) can approximate the KE-*R* results obtained in different studies by changing the parameter values. Since the equation can be fixed only by determining the parameters, it has the value of further generalization.

Table A1. Different formulas of KE-*R* relationships.

References	KE-*R* Relationships (Originally Derived)	Form	KE-*R* Relationships (Based on Equation (18))	*μ*
Carter et al., 1974 [68]	$KE = 11.32R + 0.5546R^2 - 0.5009 \times 10^{-2}R^3 + 0.126 \times 10^{-4}R^4$	Polynomial	$KE = 4.38R^{1.49}$	1.09
McGregor et al., 1976 [69]	$KE = R(27.3 + 21.68e^{-0.048R} - 41.26e^{-0.072R})$	Index	$KE = 4.86R^{1.53}$	1.30
Wischmeier et al, 1978 [70]	$KE = R(11.9 + 8.73\log R)$	Logarithm	$KE = 4.62R^{1.51}$	1.20
Bollinne et al., 1984 [71]	$KE = 12.32 \cdot R + 0.56 \cdot R^2$	Polynomial	$KE = 4.79R^{1.53}$	1.27
Steiner et al., 2000 [67]	$KE = 11R^{1.25}$	Power	$KE = 4.86R^{1.53}$	1.30
Sanchez-Moreno et al., 2012 [72]	$KE = (10.09 + 12\log R) \cdot R$	Logarithm	$KE = 4.83R^{1.53}$	1.29
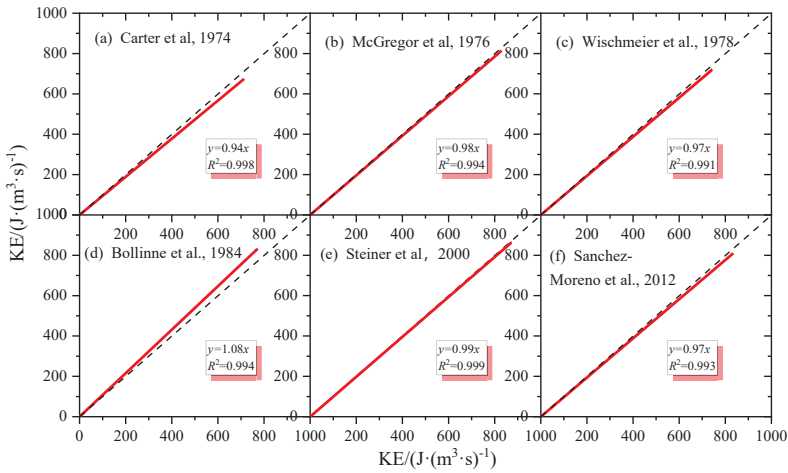
Figure A1. The approximated results of KE-*R* relationships derived by different studies are as follows: (**a**–**f**) are the linear relationships between KE-*R* obtained by this paper and KE-*R* obtained by different studies shown in Table A1, respectively.

References

1. Ding, B.; Yang, K.; Qin, J.; Wang, L.; Chen, Y.; He, X. The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. *J. Hydrol.* **2014**, *513*, 154–163. [CrossRef]
2. Chen, B.J.; Yang, J.; Pu, J.P. Statistical Characteristics of Raindrop Size Distribution in the Meiyu Season Observed in Eastern China. *J. Meteorol. Soc. Jpn.* **2013**, *91*, 215–227. [CrossRef]
3. Das, S.K.; Konwar, M.; Chakravarty, K.; Deshpande, S.M. Raindrop size distribution of different cloud types over the Western Ghats using simultaneous measurements from Micro-Rain Radar and disdrometer. *Atmos. Res.* **2017**, *186*, 72–82. [CrossRef]
4. Hazenberg, P.; Leijnse, H.; Uijlenhoet, R. The impact of reflectivity correction and accounting for raindrop size distribution variability to improve precipitation estimation by weather radar for an extreme low-land mesoscale convective system. *J. Hydrol.* **2014**, *519*, 3410–3425. [CrossRef]
5. Janapati, J.; Seela, B.K.; Reddy, M.V.; Reddy, K.K.; Lin, P.-L.; Rao, T.N.; Liu, C.-Y. A study on raindrop size distribution variability in before and after landfall precipitations of tropical cyclones observed over southern India. *J. Atmos. Sol.-Terr. Phys.* **2017**, *159*, 23–40. [CrossRef]
6. Hazenberg, P.; Yu, N.; Boudevillain, B.; Delrieu, G.; Uijlenhoet, R. Scaling of raindrop size distributions and classification of radar reflectivity–rain rate relations in intense Mediterranean precipitation. *J. Hydrol.* **2011**, *402*, 179–192. [CrossRef]
7. Waldvogel, A. TheN0Jump of Raindrop Spectra. *J. Atmos. Sci.* **1974**, *31*, 1067–1078. [CrossRef]
8. Schönhuber, M.; Randeu, W.L.; Baptista, J.P.V.P. Application of the 2D-video-distrometer for weather radar data inversion. *Phys. Chem. Earth Part B* **2000**, *25*, 1037–1042. [CrossRef]
9. Liu, X.; He, B.; Zhao, S.; Hu, S.; Liu, L. Comparative measurement of rainfall with a precipitation micro-physical characteristics sensor, a 2D video disdrometer, an OTT PARSIVEL disdrometer, and a rain gauge. *Atmos. Res.* **2019**, *229*, 100–114. [CrossRef]
10. Zhang, A.; Hu, J.; Chen, S.; Hu, D.; Liang, Z.; Huang, C.; Xiao, L.; Min, C.; Li, H. Statistical Characteristics of Raindrop Size Distribution in the Monsoon Season Observed in Southern China. *Remote Sens.* **2019**, *11*, 432. [CrossRef]
11. Raupach, T.H.; Berne, A. Correction of raindrop size distributions measured by Parsivel disdrometers, using a two-dimensional video disdrometer as a reference. *Atmos. Meas. Tech.* **2015**, *8*, 343–365. [CrossRef]
12. Bringi, V.N.; Chandrasekar, V.; Hubbert, J.; Gorgucci, E.; Randeu, W.L.; Schoenhuber, M. Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *J. Atmos. Sci.* **2003**, *60*, 354–365. [CrossRef]

13. Seliga, T.A.; Bringi, V.N. Potential use of the radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation. *J. Appl. Meteorol.* **1976**, *15*, 69–76. [CrossRef]

14. Sulochana, Y.; Rao, T.N.; Sunilkumar, K.; Chandrika, P.; Raman, M.R.; Rao, S.V.B. On the seasonal variability of raindrop size distribution and associated variations in reflectivity—Rainrate relations at Tirupati, a tropical station. *J. Atmos. Sol.-Terr. Phys.* **2016**, *147*, 98–105. [CrossRef]

15. Wen, G.; Xiao, H.; Yang, H.L.; Bi, Y.H.; Xu, W.J. Characteristics of summer and winter precipitation over northern China. *Atmos. Res.* **2017**, *197*, 390–406. [CrossRef]

16. Carollo, F.G.; Ferro, V.; Serio, M.A. Predicting rainfall erosivity by momentum and kinetic energy in Mediterranean environment. *J. Hydrol.* **2018**, *560*, 173–183. [CrossRef]

17. Ji, L.; Chen, H.N.; Li, L.; Chen, B.J.; Xiao, X.; Chen, M.; Zhang, G.F. Raindrop Size Distributions and Rain Characteristics Observed by a PARSIVEL Disdrometer in Beijing, Northern China. *Remote Sens.* **2019**, *11*, 1479. [CrossRef]

18. Angulo-Martínez, M.; Barros, A.P. Measurement uncertainty in rainfall kinetic energy and intensity relationships for soil erosion studies: An evaluation using PARSIVEL disdrometers in the Southern Appalachian Mountains. *Geomorphology* **2015**, *228*, 28–40. [CrossRef]

19. Carollo, F.G.; Serio, M.A.; Ferro, V.; Cerdà, A. Characterizing rainfall erosivity by kinetic power - Median volume diameter relationship. *Catena* **2018**, *165*, 12–21. [CrossRef]

20. Meshesha, D.T.; Tsunekawa, A.; Tsubo, M.; Haregeweyn, N.; Tegegne, F. Evaluation of kinetic energy and erosivity potential of simulated rainfall using Laser Precipitation Monitor. *Catena* **2016**, *137*, 237–243. [CrossRef]

21. Abd Elbasit, M.A.M.; Yasuda, H.; Salmi, A.; Anyoji, H. Characterization of rainfall generated by dripper-type rainfall simulator using piezoelectric transducers and its impact on splash soil erosion. *Earth Surf. Process. Landf.* **2010**, *35*, 466–475. [CrossRef]

22. Serio, M.A.; Carollo, F.G.; Ferro, V. Raindrop size distribution and terminal velocity for rainfall erosivity studies. A review. *J. Hydrol.* **2019**, *576*, 210–228. [CrossRef]

23. Fan, Y.; Chen, Y.; Wei, J. An Analysis of Drought Features in Shaanxi Province. *J. Xi'an Univ. Technol.* **1996**, *12*, 200–206.

24. Li, Y.; Cao, Z.; Long, H.; Liu, Y.; Li, W. Dynamic analysis of ecological environment combined with land cover and NDVI changes and implications for sustainable urban–rural development: The case of Mu Us Sandy Land, China. *J. Clean. Prod.* **2017**, *142*, 697–715. [CrossRef]

25. Xie, Z.; Yang, H.; Lv, H. Study on the relationship between rainfall kinetic energy and rainfall intensity based on raindrop spectrum observations. *Water Resour. Hydropower Eng.* **2019**. under review.

26. Ulbrich, C.W. Natural Variations in the Analytical Form of the Raindrop Size Distribution. *J. Clim. Appl. Meteorol.* **1983**, *22*, 1764–1775. [CrossRef]

27. Testud, J.; Oury, S.; Black, R.A.; Amayenc, P.; Dou, X.K. The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *J. Appl. Meteorol.* **2001**, *40*, 1118–1140. [CrossRef]

28. Ulbrich, C.W.; Atlas, D. Rainfall microphysics and radar properties: Analysis methods for drop size spectra. *J. Appl. Meteorol.* **1998**, *37*, 912–923. [CrossRef]

29. Tokay, A.; Wolff, D.B.; Petersen, W.A. Evaluation of the New Version of the Laser-Optical Disdrometer, OTT Parsivel2. *J. Atmos. Ocean. Technol.* **2014**, *31*, 1276–1288. [CrossRef]

30. Zhang, H.; HE, H.; Zhang, Y.; Zeng, Q.; Bai, S. Parameter Characteristics Analysis of Raindrop Spectrum Fitting Models in Nanjing. *Meteorol. Environ. Sci.* **2017**, *40*, 77–84.

31. Seela, B.K.; Janapati, J.; Lin, P.L.; Wang, P.K.; Lee, M.T. Raindrop Size Distribution Characteristics of Summer and Winter Season Rainfall Over North Taiwan. *J. Geophys. Res.-Atmos.* **2018**, *123*, 11602–11624. [CrossRef]

32. Zhang, G.; Vivekanandan, J.; Brandes, E.A. The Shape–Slope Relation in Observed Gamma Raindrop Size Distributions: Statistical Error or Useful Information? *J. Atmos. Ocean. Technol.* **2003**, *20*, 1106–1119. [CrossRef]

33. Low, T.B.; List, R. Collision, Coalescence and Breakup of Raindrops. Part I: Experimentally Established Coalescence Efficiencies and Fragment Size Distributions in Breakup. *J. Atmos. Sci.* **1982**, *39*, 1591–1606. [CrossRef]

34. Tokay, A.; Bashor, P.G. An Experimental Study of Small-Scale Variability of Raindrop Size Distribution. *J. Appl. Meteorol. Clim.* **2010**, *49*, 2348–2365. [CrossRef]

35. Zhou, L.M.; Zhang, H.S.; Wang, J.; Wang, Q.; Chen, X.L. Raindrop Spectral Charcteristics of Mixed-cloud Precipitation in Shandong Province. *Meteorol. Sci. Technol.* **2010**, *38*, 73–77.

36. Li, H.; Yin, Y.; Shan, Y.; Jin, Q. Statistical Characteristics of Raindrop Size Distribution for Stratiform and Convective Precipitation at Different Altitudes in Mt. Huangshan. *Chin. J. Atmos. Sci.* **2018**, *42*, 268–280.

37. Chen, C.; Yin, Y.; Chen, B.J. Raindrop Size Distribution at Different Altitudes in Mt. Huang. *Trans. Atmos. Sci.* **2015**, *38*, 388–395. [CrossRef]

38. Kinnell, P. Rainfall Intensity-Kinetic Energy Relationships for Soil Loss Prediction 1. *Soil Sci. Soc. Am. J.* **1981**, *45*, 153–155. [CrossRef]

39. Atlas, D.; Ulbrich, C.W. Path-and area-integrated rainfall measurement by microwave attenuation in the 1–3 cm band. *J. Appl. Meteorol.* **1977**, *16*, 1322–1331. [CrossRef]

40. Marshall, J.S.; Palmer, W.M.K. The Distribution of Raindrops with Size. *J. Meteorol.* **1948**, *5*, 165–166. [CrossRef]

41. Uijlenhoet, R.; Stricker, J. A consistent rainfall parameterization based on the exponential raindrop size distribution. *J. Hydrol.* **1999**, *218*, 101–127. [CrossRef]

42. Lane, J.; Hackathorn, M.; Kewley, J.; Madore, M.; May, M.; Briggs, C.; DeLeon Springs, F. A method for estimating 3-D spatial variations of rainfall drop size distributions over remote ocean areas. Presented at the Fifth International Conference on Remote Sensing for Marine and Coastal Environments, San Diego, CA, USA, 5–7 October 1998; p. 7.

43. Ochou, A.D.; Nzeukou, A.; Sauvageot, H. Parametrization of drop size distribution with rain rate. *Atmos. Res.* **2007**, *84*, 58–66. [CrossRef]

44. Li, P.; Sun, X.; Zhao, X. Analysis of precipitation and potential evapotranspiration in arid and semi-arid area of China in recent 50 years. *J. Arid Land Res. Environ.* **2012**, *7*, 57–63.

45. Beard, K.V. Terminal Velocity and Shape of Cloud and Precipitation Drops Aloft. *J. Atmos. Sci.* **1976**, *33*, 851–864. [CrossRef]

46. Hunt, W.F.; Jarrett, A.R.; Smith, J.T.; Sharkey, L.J. Evaluating bioretention hydrology and nutrient removal at three field sites in North Carolina. *J. Irrig. Drain. Eng.* **2006**, *132*, 600–608. [CrossRef]

47. Hunter, S.M. WSR-88D radar rainfall estimation: Capabilities, limitations and potential improvements. *Natl. Weather Dig.* **1996**, *20*, 26–38.

48. Austin, P.M. Relation between Measured Radar Reflectivity and Surface Rainfall. *Mon. Weather Rev.* **1987**, *115*, 1053–1071. [CrossRef]

49. Sauvageot, H. Rainfall Measurement by Radar—A Review. *Atmos. Res.* **1994**, *35*, 27–54. [CrossRef]

50. Wilson, J.W.; Brandes, E.A. Radar measurement of rainfall—A summary. *Bull. Am. Meteorol. Soc.* **1979**, *60*, 1048–1060. [CrossRef]

51. Zhang, P.; Du, B.; Dai, T.P. *Radar Meteorology*; China Meteorological Press: Beijing, China, 2001.

52. Petan, S.; Rusjan, S.; Vidmar, A.; Mikos, M. The rainfall kinetic energy-intensity relationship for rainfall erosivity estimation in the mediterranean part of Slovenia. *J. Hydrol.* **2010**, *391*, 314–321. [CrossRef]

53. Battaglia, A.; Rustemeier, E.; Tokay, A.; Blahak, U.; Simmer, C. PARSIVEL Snow Observations: A Critical Assessment. *J. Atmos. Ocean. Technol.* **2010**, *27*, 333–344. [CrossRef]

54. Liu, X.C.; Gao, T.C.; Liu, L. A comparison of rainfall measurements from multiple instruments. *Atmos. Meas. Tech.* **2013**, *6*, 1585–1595. [CrossRef]

55. Kang, L. *A Study on Dynamic Correction for Radar-Derived Quantitation Precipitation Estimation Based on the Optimal Z-I Relationship*; Lanzhou University: Lanzhou, China, 2014.

56. Yuan, X.; Ni, G.; Pan, A.; Wei, L. NEXRAD Z-R Power Relationship in Beijing Based on Optimization Algorithm. *J. Chin. Hydrol.* **2010**, *30*, 1–6.

57. Zhang, P.; Dai, T.; WANG, D.; Lin, B. Derivation of The Z-I Relationship by Optimization and The Accuracy in The Quantitative Rainfall Measurement. *Sci. Meteorol. Sin.* **1992**, *12*, 333–338.

58. Marshall, J.S.; Langille, R.C.; Palmer, W.M.K. Measurement of Rainfall by Radar. *J. Meteorol.* **1947**, *4*, 186–192. [CrossRef]

59. Chen, Q.; Niu, S.; Zhang, Y.; Xu, F. Z-R Relationship from the Particle Size and Velocity (Parsivel) optical disdrometer and its Application in Estimating Areal Rainfall. In Proceedings of the 2008 2nd International Conference on Bioinformatics and Biomedical Engineering, Shanghai, China, 16–18 May 2008.

60. Marzuki; Randeu, W.L.; Kozu, T.; Shimomai, T.; Hashiguchi, H.; Schönhuber, M. Raindrop axis ratios, fall velocities and size distribution over Sumatra from 2D-Video Disdrometer measurement. *Atmos. Res.* **2013**, *119*, 23–37. [CrossRef]

61. Blanchard, D.C. Raindrop Size-Distribution in Hawaiian Rains. *J. Meteorol.* **1953**, *10*, 457–473. [CrossRef]

62. Sivaramakrishnan, M.V. Studies of raindrop size characteristics in different types of tropical rain using a simple recorder. *Indian J. Meteorol. Geophys.* **1961**, *12*, 189–217.

63. Yakubu, M.L.; Yusop, Z.; Fulazzaky, M.A. The influence of rain intensity on raindrop diameter and the kinetics of tropical rainfall: Case study of Skudai, Malaysia. *Hydrolog. Sci. J.* **2016**, *61*, 944–951. [CrossRef]

64. Takeuchi, D. Characterization of raindrop size distributions. In Proceedings of the Conference on Cloud Physics and Atmospheric Electricity, Issaquah, WA, USA, 31 July–4 August 1978; pp. 154–161.

65. Teng, X. *Sensitivity of the W Band Airborne Cloud Radar Reflectivity Factor Z to Cloud Parameters*; Nanjing University of Information Science and Technology: Nanjing, China, 2011.

66. Park, S.W.; Mitchell, J.K.; Bubenzer, G.D. An analysis of splash erosion mechanics. In Proceedings of the ASAE 1980 Winter Meeting, Chicago, IL, USA, 4 December 1980; p. 27.

67. Steiner, M.; Smith, J.A. Reflectivity, rain rate, and kinetic energy flux relationships based on raindrop spectra. *J. Appl. Meteorol.* **2000**, *39*, 1923–1940. [CrossRef]

68. Carter, C.E.; Greer, J.D.; Braud, H.J.; Floyd, J.M. Raindrop Characteristics in South Central United-States. *Trans. ASAE* **1974**, *17*, 1033–1037. [CrossRef]

69. McGregor, K.C.; Mutchler, C.K. Status of the R factor in northern Mississippi. In *Soil Erosion: Prediction and Control, Soil Conservation Soc. Amer.*; Nature: UK, 1976; pp. 135–142.

70. Wischmeier, W.H.; Smith, D.D. *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning*; Science and Education Administration, US Department of Agriculture: Washington, WA, USA, 1978; Volume 537.

71. Bollinne, A.; Florins, P.; Hecq, P.; Homerin, D.; Renard, V. Etude de l'énergie des pluies en climat tempéré océanique d'Europe Atlantique. *Zeitschrift für Geomorphologie* **1984**, *49*, 27–35.

72. Sanchez-Moreno, J.F.; Mannaerts, C.M.; Jetten, V.; Loffler-Mang, M. Rainfall kinetic energy-intensity and rainfall momentum-intensity relationships for Cape Verde. *J. Hydrol.* **2012**, *454*, 131–140. [CrossRef]

MDPI

Article

A Preliminary Assessment of the Gauge-Adjusted Near-Real-Time GSMaP Precipitation Estimate over Mainland China

Dekai Lu [1,2] and Bin Yong [1,2,*]

[1] State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China; sdludekai@hhu.edu.cn

[2] School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China

*   Correspondence: yongbin@hhu.edu.cn; Tel.: +86-25-8378-7485

Abstract: The near-real-time satellite-derived precipitation estimates are attractive for a wide range of applications like extreme precipitation monitoring and natural hazard warning. Recently, a gauge-adjusted near-real-time GSMaP precipitation estimate (GSMaP_Gauge_NRT) was produced to improve the quality of the original GSMaP_NRT. In this study, efforts were taken to investigate and validate the performance of the GSMaP_Gauge_NRT using gauge observations over Mainland China. The analyses indicated that GSMaP_NRT generally overestimated the gauge precipitation in China. After calibration, the GSMaP_Gauge_NRT effectively reduced this bias and was more consistent with gauge observations. Results also showed that the correction scheme of GSMaP_Gauge_NRT mainly acted on hit events and could hardly make up the miss events of the satellite precipitation estimates. Finally, we extended the evaluation to the global scale for a broader view of GSMaP_Gauge_NRT. The global comparisons exhibited that the GSMaP_Gauge_NRT was in good agreement with the GSMaP_Gauge product. In conclusion, the GSMaP_Gauge_NRT had better performance than the GSMaP_NRT and was a more reliable near-real-time satellite precipitation product.

Keywords: satellite precipitation; Mainland China; GSMaP_NRT; GSMaP_Gauge_NRT

1. Introduction

Reliable precipitation estimates are crucial because of their role in flood monitoring, crop yield, and water resource management [1–3]. However, in many regions of Earth, like the oceans, deserts, and mountains, ground-based observing networks from gauges and radars are sparse or even nonexistent, which restricts our understanding of global water cycle and local hydrological processes [4–6]. The recent development of precipitation-retrieval techniques from satellite-based remote sensing makes it possible of measuring precipitation on the global scale. The remote sensing of precipitation combines the advantage of the frequency sampling of infrared (IR) sensors derived from geostationary (GEO) satellites and the superior accuracy (but poor sampling) of passive microwave (PMW) sensors carried onboard the low earth orbiting (LEO) satellites, in an effort to produce precipitation data with extensive spatial coverage and fine resolutions [7–9].

To date, various satellite precipitation missions have been implemented and their products have been made available to the public. Previous satellite precipitation missions include the NASA's Tropical Rainfall Measuring Mission (TRMM [10]), NOAA's Climate Prediction Center (CPC) morphing technique (CMORPH [11]), JAXA's Global Satellite Mapping of Precipitation (GSMaP [12]), the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN [13]), the Climate Hazard Group InfraRed Precipitation (CHIRP [14]), and the successor of TRMM: Global Precipitation Measurement (GPM [3]). These satellite precipitation missions and

products have benefitted the hydrology and meteorology community in relevant researches and applications. However, more recent studies have found that satellite precipitation products contain considerable errors due to the indirect retrieval methods of remote sensing [15–23]. One of the most effective strategies to diminish these errors is combining ground-based data to adjust the satellite precipitation products. For example, the 3B42 product is one of the most popular gauge-adjusted satellite precipitation products at the TRMM era, which incorporates the monthly Global Precipitation Climatology Centre (GPCC) gauge data to increase the accuracy of its original satellite-only 3B42RT product. Today, numerous studies have demonstrated that the 3B42 shows substantial improvement than the 3B42RT, with lower bias and better detection skills [24–28]. However, the 3B42 data is not available in real-time, and researchers must wait ~2.5 months after observation time, while the delay of pure satellite-derived 3B42RT is only 8 h [29]. Obviously, traditional gauge-adjusted schemes depend on the time availability of gauge data. Nevertheless, collecting and processing gauge data in real-time is not possible on a global scale, especially in underdeveloped countries and areas, which delay the availability of satellite precipitation products. The question of the availability exists widely in the gauge-adjusted satellite precipitation products. For example, if someone want to use PERSIANN-CDR or GSMaP_Gauge data, they must wait ~3 months or ~3 days after observation, respectively [21,30].

In some cases, the real-time availability precipitation data is more critical for applications like rainstorm monitoring and flash flood warning, and it does not seem practical to use the traditionally delayed gauge-adjusted satellite precipitation products [31,32]. Thus, it is important to reduce the error of satellite precipitation estimates as much as possible without jeopardizing its near-real-time availability. To this end, a climatological calibration algorithm (CCA) was proposed in the TRMM Multisatellite Precipitation Analysis (TMPA) real-time system. This method utilizes climatological gauge information to alleviate errors and keep the timeliness of 3B42RT itself [9]. Yong et al. [31] initially investigated the performance of CCA in the 3B42RT precipitation estimates over two different basins of China using a local dense rain gauge. The author found that the systematic errors in 3B42RT were minimized overall after the CCA calibration. Nevertheless, the author also highlighted that the performance of calibrated precipitation became worse in high-latitude areas, or areas beyond the 40° latitude belts. In addition, from a global map view of error analysis, Yong et al. [9] demonstrated that the CCA calibrated 3B42RT precipitation has large bias in mountainous regions (especially over the Tibetan Plateau). Theoretically, the CCA is used in the GPM near-real-time runs of the Integrated Multisatellite Retrievals for GPM (IMERG) algorithm. However, considering the unstable performance of CCA, the developers of the IMERG algorithm are re-evaluating the CCA calibration. Meantime, the IMERG near-real-time products do not currently have climatological calibration.

As the Japanese counterpart of IMERG, the GSMaP is another mainstream satellite precipitation product at GPM era, which was produced by reliable physical models and by distributing hourly global precipitation map with $0.1° \times 0.1°$ resolution [20,33,34]. To satisfy different application requirements, there are two main groups of GSMaP products: Near-real-time and standard products. As the name implies, the near-real-time product is intended to provide available satellite precipitation quickly, while the standard product applies more PMW/IR sources to create relative accurate precipitation estimates. Correspondingly, the near-real-time product has about a 3-h delay, and the standard product has a large latency of about 3 d. To reduce bias on the satellite-derived GSMaP products, gauge-adjusted GSMaP products are developed using ground-gauge measurement as a calibrator. The gauge-calibrated product of standard GSMaP_MVK is GSMaP_Gauge, which adjusted by daily CPC gauge data. Many studies have assessed and compared the performances of GSMaP_Gauge over the last few years and have shown that GSMaP_Gauge is a satisfactory gauge-adjusted satellite precipitation estimation around the world, especially over East Asia [34–38].

Recently, in the GSMaP project, the GSMaP_Gauge_NRT was produced by a GSMaP algorithm team, aiming to improve the accuracy of near-real-time product of GSMaP (i.e., GSMaP_NRT) and maintain its timeliness. Section 2.2 describes the calibration procedure of GSMaP_Gauge_NRT in detail. Thus, it is crucial to understand the performance of new GSMaP_Gauge_NRT product timely. In the

official document of GSMaP, the GSMaP developers eagerly encouraged people to evaluate and validate
the GSMaP_Gauge_NRT in different regions. Therefore, in this study, we systematically assessed
the performance of the GSMaP_Gauge_NRT precipitation estimates and the original uncalibrated
GSMaP_NRT over the Mainland China. The rest of this paper is organized as follows. In Section 2, we
describe the study area, the precipitation data and the error metrics. Then, a presentation of results
and discussion in this study are provided in Sections 3 and 4, respectively. Finally, the summary and
conclusions are given in Section 5.

2. Materials and Methods

2.1. Study Area

China, which covers an area of about 9.6 million km^2, is located in the eastern Asia. The terrain of
China is complex and varied, with flat plains and hills in the east and high mountains and plateaus in
the west. The elevation of China has an overall descending tendency from west to the east. Specially,
the world's "Third Pole", the Tibetan Plateau (TP), is situated in southwest China. Generally, the climate
on the China features hot and rainy summers and cold and dry winters due to the effects of monsoon
system and topography. In addition, the annual precipitation in China varies greatly, with an obvious
gradient from southeast to northwest, except for some parts of northwestern Xinjiang. Over southeast
China, a huge amount of water vapor from the Pacific Ocean is brought by the Asian monsoon in the
summer, producing abundant precipitation (exceeding 800 mm/year). Over the inland areas of western
China, which are far from the ocean, monsoon precipitation becomes negligible and the continental
climate dominates (with annual precipitation less than 400 mm). Considering the diverse climate in
China, following Yong et al. [39], we separated mainland China into four representative climate regions
based on the annual precipitation (Figure 1a): (1) The humid region, mainly covering the southeast
of China and dominated by the subtropical monsoon with the average annual precipitation (AAP)
above 800 mm; (2) the semi-humid region (AAP between 400–800 mm), including northeast China,
North China, and southeast parts of TP; (3) the semi-arid region (AAP of 200–400 mm), which consists
of two parts, one region extending from the southwest TP into the North Inner Mongolia with a narrow
strip, and another region located in the Ili river valley; (4) the arid region (AAP < 200 mm), a main
desert areas in China, including the most of Xinjiang and western part of Inner Mongolia.
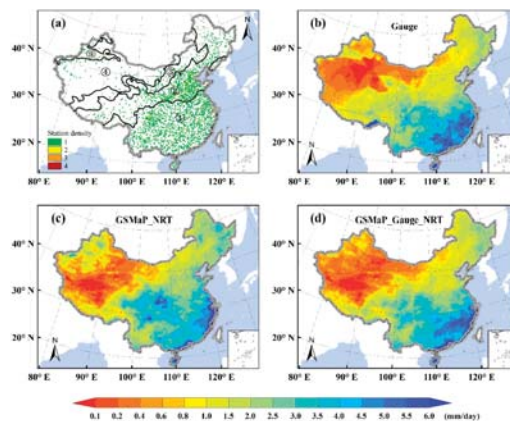
Figure 1. (**a**) Distribution of station density used in this study. Four climate regions (①–④: Humid,
semi-humid, semi-arid, arid) are outlined in Mainland China. (**b**–**d**) Daily mean precipitation
over Mainland China during the period of September 2017–August 2019, from gauge observations,
GSMaP_NRT, and GSMaP_Gauge_NRT precipitations.

2.2. Data Sources

2.2.1. Rain Gauge Data

A gridded precipitation product (0.25° × 0.25°) of the China Gauge-Based Daily Precipitation Analysis (CGDPA) was used as the ground reference dataset. This dataset was produced and is routinely operated by the National Meteorological Information Center (NMIC) of the China Meteorological Administration (CMA). Based on amassing ~2400 gauge stations across Mainland China, the NMIC uses a modified interpolation method of climatology-based optimal interpolation (OI) proposed by Xie et al. [40] to yield the daily gridded CGDPA product. All the gauge data inputted in the CGDPA are manually recorded by bucket rain gauges and must pass a strict quality control procedure including the extreme values' check, internal consistency check, and spatial consistency check [41]. The validation reports of Shen and Xiong [42] demonstrated that the CGDPA is a more reliable precipitation product than other existing surface observation datasets over Mainland China, and the authors recommended it as a reliable ground-based data to various meteorological and hydrological applications. At present, the CGDPA is available as the official gridded daily precipitation dataset in the CMA website and can be downloaded from http://data.cma.cn/site/index.html. Considering the ease of access and high-quality of CGDPA, a large number of studies have used the CGDPA to assess the satellite precipitation products over China [39,43–45]. Figure 1a shows the number of gauge stations in each 0.25° grid box of CGDPA. Clearly, about 80% of the gauges are located in the eastern China with low elevations, while relatively sparse gauges are distributed over the western mountains and deserts, especially over the TP.

2.2.2. GSMaP Near-Real-Time Precipitation Products

As one of Japanese GPM projects, GSMaP was implemented under the Japanese Precipitation Measuring Mission (PMM) science team with the target of providing a global precipitation map with high precision and high resolution [34]. The GSMaP algorithm uses various of PMW/IR sensors to produce the "best" precipitation estimates through several steps. First, several PMW radiometers carried by different satellites, such as the GPM microwave imager (GMI), TRMM microwave imager (TMI), special sensor microwave imager/sounder (SSMIS), advanced microwave scanning radiometer 2 (AMSR2), advanced microwave sounding unit-A (AMSU-A), and microwave humidity sounder (MHS), are used to retrieve quantitative precipitation estimates [33,46]. Then, it uses the cloud motion vector derived from successive geo-IR images to propagate the precipitation area for filling the gaps between PMW-based estimates, which is similar to CMORPH. In addition, a Kalman filter model is applied to modify precipitation rates after propagation. Finally, the forward and backward propagated precipitation estimates are weighted and combined to produce the standard GSMaP_MVK product. At the beginning of design, the GSMaP algorithm developers did not consider near-real-time operation and data availability. To meet this demand, a near-real-time product of the GSMaP (GSMaP_NRT) with resolutions of 0.1° and 1 h was developed. Different from GSMaP_MVK, the GSMaP_NRT only employs forward cloud movement to hold operability in near-real-time. The emergence of GSMaP_NRT product attracts a lot of data users, owing to its short latency (~3 h after observation). To reduce the bias in the GSMaP_NRT product, a new algorithm introducing gauge information to GSMaP_NRT (i.e., GSMaP_Gauge_NRT) is currently under development. In the GSMaP_Gauge_NRT product, a precipitation error parameters model was created based on the historical database of GSMaP_Gauge. Then, these parameters were used to adjust the GSMaP_NRT estimation in near-real-time to improve the precision of GSMaP_NRT. Considering that the GSMaP_Gauge_NRT product does not use the gauge measurement directly, this makes the GSMaP_Gauge_NRT independent of the ground gauge observations. Recently, the latest GSMaP algorithm upgraded to version 7, and its near-real-time products were made available after 17 January 2017. In this study, the GSMaP_Gauge_NRT product and the uncalibrated GSMaP_NRT product were investigated over a complete two-year period (from September 2017 to August 2019). Both satellite precipitations were aggregated into daily amounts, with a 0.25° × 0.25° resolution corresponding to the gauge data.

2.3. Methods

In this study, we conducted the assessment and comparison of GSMaP precipitation based on continuous statistical metrics and contingency table metrics. The continuous metrics included correlation coefficient (CC), mean error (ME), root mean squared error (RMSE), and relative bias (BIAS), which are widely used to quantitatively represent the degree of agreement and the error between satellite precipitation and gauge observations. These continuous metrics were calculated by the following equation:

$$CC = \frac{\sum_{i=1}^{n}(G_i - \overline{G})(S_i - \overline{S})}{\sqrt{\sum_{i=1}^{n}(G_i - \overline{G})^2} \cdot \sqrt{\sum_{i=1}^{n}(S_i - \overline{S})^2}} \tag{1}$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}(S_i - G_i) \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - G_i)^2} \tag{3}$$

$$BIAS = \frac{\sum_{i=1}^{n}(S_i - G_i)}{\sum_{i=1}^{n}G_i} \times 100\% \tag{4}$$

where S_i and G_i are the precipitation values from satellite estimation and gauge data, respectively; correspondingly, $\overline{S}$ and $\overline{G}$ are their mean precipitation, and n is the number of samples.

In addition, three contingency table metrics were adopted to evaluate the capability of satellite precipitation in the detection of precipitation events. These categorical metrics were the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). POD is usually used to represent the fraction of precipitation events that correctly detected by satellite among all the actual precipitation events. FAR denotes the ratio of false alarm by satellite among the total satellite detected events. The CSI, combining the correct hit, false alarm, and missed event, is a more comprehensive score. The formulas of these contingency table metrics are listed below:

$$POD = \frac{H}{H + M} \tag{5}$$

$$FAR = \frac{F}{H + F} \tag{6}$$

$$CSI = \frac{H}{H + M + F} \tag{7}$$

where H, M, and F are the numbers of different precipitation events: Hit (both satellite estimates and gauge observations detect rain), miss (observed rain that is not detected by satellite), false (rain detected but not observed). Here, a commonly used threshold of 1.0 mm/day was set to define the rain/no rain event, as suggested by many previous studies [47–50].

For more detailed description of above continuous statistical metrics and contingency table metrics, readers can refer to Yong et al. [25] and Lu et al. [30]. We need to point out that all metrics were calculated in the 0.25° × 0.25° grid boxes with at least one gauge in order to ensure more convincing results (gauge distribution shows in Figure 1a). However, we also calculated metric value in every grid box to enable a visualization when presenting continuous spatial distribution (Figure 2).
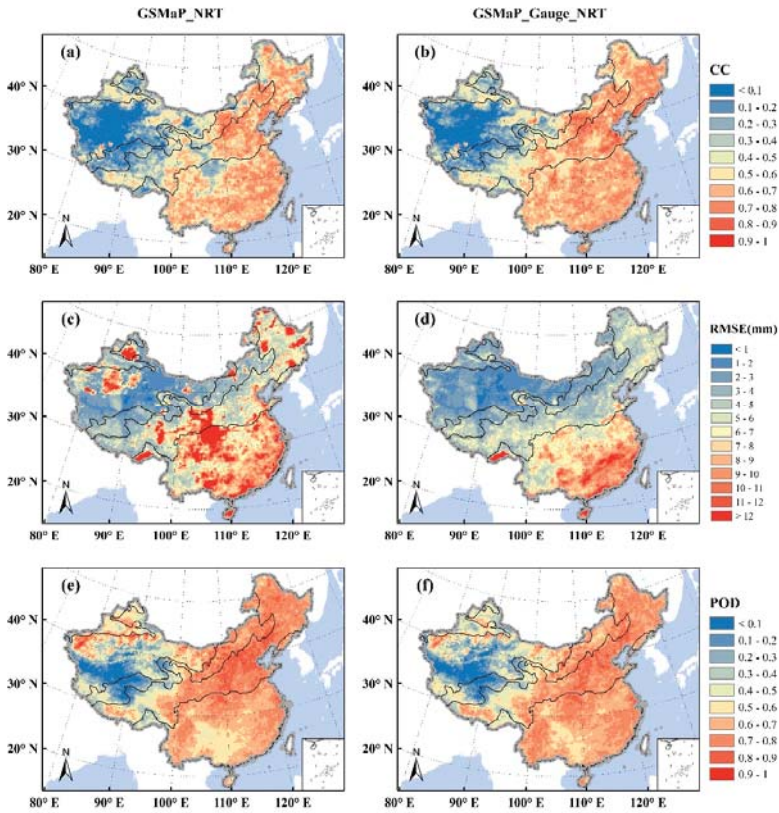
Figure 2. Spatial distribution of statistical indices derived from the GSMaP_NRT (left column) and GSMaP_Gauge_NRT (right column) daily precipitation at 0.25° × 0.25° resolution over the Mainland China: (**a**,**b**) Correlation coefficient (CC), (**c**,**d**) root mean square error (RMSE), and (**e**,**f**) probablility of detection (POD).

3. Results

3.1. Daily Mean Precipitation

Figure 1b–d displays the spatial distributions of two-year daily mean precipitation for GSMaP and CGDPA precipitation products. Generally speaking, the spatial distributions of GSMaP and CGDPA precipitations were similar, showing a downward gradient from the southeast China to the northwest China. However, a pronounced difference in the precipitation amount was found between the gauge observations and GSMaP satellite precipitation products. For example, compared with the CGDPA, the GSMaP_NRT tended to underestimate the gauge observations in the southeast and overestimated them in the northwest. Impressively, the GSMaP_NRT significantly overestimated the gauge precipitation in the Sichuan province due to the indirect retrieval of satellite precipitation. After the parameterized gauge calibration, the errors in GSMaP_NRT were effectively suppressed, and the GSMaP_Gauge_NRT had a more reliable performance than the GSMaP_NRT in capturing the spatial patterns of precipitation over China. Therefore, the GSMaP_Gauge_NRT combined historical gauge information to reduce biases, making it consistent with the ground measurements. This suggests that the parameterized adjustment procedures can effectively enhance the quality of the original GSMaP_NRT satellite precipitation estimates.

3.2. Comparison and Validation of GSMaP_NRT and GSMaP_Gauge_NRT Products

Figure 2 shows the spatial maps of CC, RMSE, and POD, which were computed from two GSMaP products against gauge observations over the Mainland China at the 0.25° × 0.25° resolution grid. Generally speaking, the spatial distributions of GSMaP_Gauge_NRT exhibited a great improvement compared to that of the GSMaP_NRT, with higher CC, lower RMSE, and slightly better POD values. Over Mainland China, the CC increased from 0.58 with GSMaP_NRT to 0.67 with GSMaP_Gauge_NRT and the RMSE dropped from 9.11 mm to 7.07, besides a small change of POD between GSMaP_NRT (0.69) and GSMaP_Gauge_NRT (0.70). With respect to the spatial performance of error metrics, both GSMaP_NRT and GSMaP_Gauge_NRT products exhibited similar features. That is, worse values of CC and POD occurred in the northwest and improved toward the southeast, while higher RMSE existed in the southeast. This phenomenon was reasonable because the RMSE value increased with increasing precipitation amounts, and southeastern China has more precipitation than other areas in China. Distributions of CC, RMSE, and POD indicate that GSMaP_Gauge_NRT performs better than GSMaP_NRT over Mainland China, suggesting that the calibration in near-real-time can effectively reduce the error and improve detectability of GSMaP_NRT.

We further inquired about the temporal behavior and the seasonal statistics of the GSMaP_NRT and GSMaP_Gauge_NRT products over Mainland China. In order to ensure a more accurate comparison, only those grids that contained at least one gauge were taken to compute the statistical indices. Figure 3 depicts the monthly precipitation and monthly variations of statistical metrics from gauges and GSMaP precipitation products by calculating at daily scale. Table 1 summarizes the seasonal statistics including spring (March–May), summer (June–August), autumn (September–November), and winter (December–February). It can be see that both GSMaP_NRT and GSMaP_Gauge_NRT products can generally capture the intra-annual and seasonal variation patterns of precipitation over China, with the rainy summer and dry winter (Figure 3a). The GSMaP_NRT showed much more precipitation than gauge observations in most months, and the GSMaP_Gauge_NRT reduced this overestimation, which was more consistent with gauge observations. The time series of statistical indices clearly show that the GSMaP_Gauge_NRT outperforms GSMaP_NRT with higher correlation, lower error, and better detection (Figure 3b–d). This further confirms that the calibration in the GSMaP_Gauge_NRT can substantially improve the quality of the original GSMaP_NRT precipitation product. However, it is worth noting that the BIAS was increased in some months. This issue may be due to the fact that the overestimation and underestimation at different regions could cancel each other out when calculating the BIAS value. Focusing on the curves of CC and RMSE, we can conclude that the GSMaP_Gauge_NRT had better agreement with gauge observations than the GSMaP_NRT products over China. In addition, we note that the performance of satellite precipitation productions showed obvious seasonally dependent variations, with better statistical indices in summer and worse in winter (Table 1). Taking CC as an example, the value of CC decreased from 0.62 in summer to 0.41 in winter for GSMaP_NRT and from 0.67 to 0.58 for GSMaP_Gauge_NRT. During the winter months, the snow brought by the westerly winds was the main form of precipitation in the north and west regions of China. However, the complex radiative properties of ice particles and snowflakes restricted the retrieval capability of microwave radiation. On the one hand, the low-frequency channels of PMW sensors contain limited snow detection information, but account for most of PMW channels and are traditionally used to retrieve rain drops. On the other hand, more snow-related high-frequency channels will be seriously interfered in the snow-covered background surfaces, which often present a similar passive microwave signature as the falling snow [51]. Therefore, measuring solid precipitation and snow is a challenging task for the satellite precipitation retrievals, which was pointed by many previous studies [15,23,25,52]. The GSMaP_Gauge_NRT product can combine the gauge information to reduce the precipitation error of GSMaP_NRT in winter, showing more consistency with gauge observations.
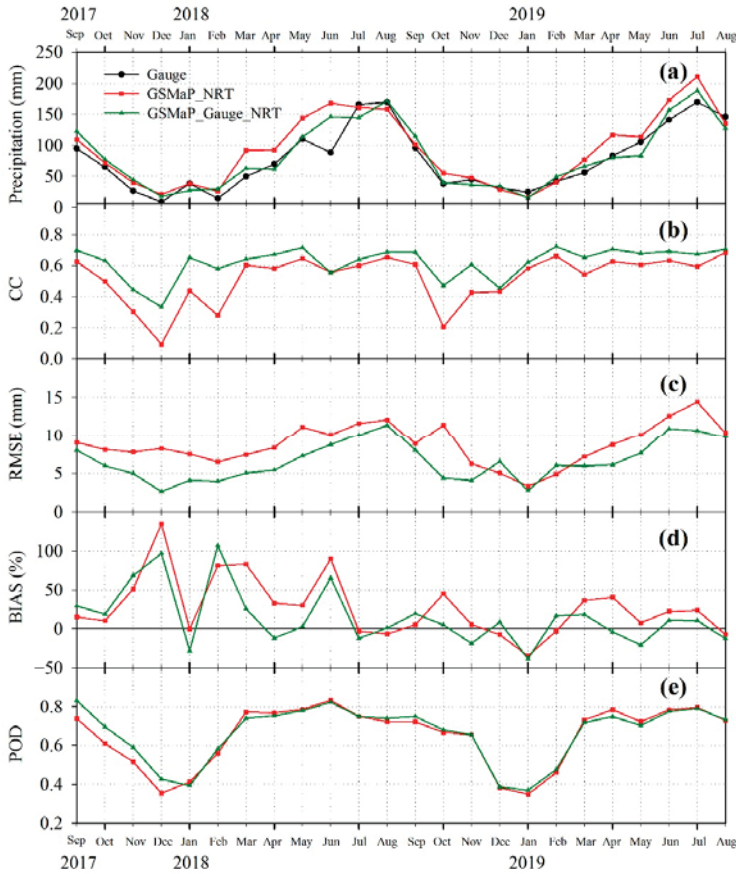
Figure 3. (**a**) Time series of mean monthly precipitation and monthly variations of statistical indices over grid boxes with at least one gauge in Mainland China: (**b**) Correlation coefficient (CC), (**c**) root mean square error (RMSE), (**d**) relative bias (BIAS), and (**e**) probablility of detection (POD).

Table 1. Seasonal statistics of GSMaP_NRT and GSMaP_Gauge_NRT against ground observations from selected 0.25° grid boxes over the Mainland China.

Season	Product	CC	ME (mm)	BIAS (%)	RMSE (mm)	POD	FAR	CSI
Annual	GSMaP_NRT	0.58	0.40	15.84	9.11	0.69	0.33	0.52
	GSMaP_Gauge_NRT	0.67	0.11	4.46	7.07	0.70	0.35	0.51
Spring	GSMaP_NRT	0.61	0.88	34.49	8.91	0.76	0.36	0.54
	GSMaP_Gauge_NRT	0.68	−0.03	−1.23	6.29	0.74	0.34	0.53
Summer	GSMaP_NRT	0.62	0.33	6.48	12.26	0.76	0.26	0.60
	GSMaP_Gauge_NRT	0.67	0.01	0.27	10.59	0.76	0.24	0.61
Autumn	GSMaP_NRT	0.48	0.32	16.15	8.71	0.67	0.36	0.48
	GSMaP_Gauge_NRT	0.65	0.39	19.41	6.12	0.72	0.45	0.45
Winter	GSMaP_NRT	0.41	0.06	7.07	6.15	0.42	0.41	0.32
	GSMaP_Gauge_NRT	0.58	0.07	8.33	4.56	0.43	0.45	0.32

Considering the diverse climate of China, it was rational to subdivide national-scale evaluation into regional analyses. Figure 4 shows the scatterplots of daily GSMaP_NRT and GSMaP_Gauge_NRT against gauge observations for the selected grids over different climate regions. Clearly, over the four

climate regions, all the scatterplots show that the scatter points of GSMaP_Gauge_NRT were clustered closer to the 1:1 line than those of GSMaP_NRT estimates, meaning that the GSMaP_Gauge_NRT was more in agreement with gauge observations. The GSMaP_NRT estimate significantly overestimated the gauge precipitation with BIAS range from 0.82% to 149.08% (see left column in Figure 4). After the calibration, these biases were effectively minimized in the GSMaP_Gauge_NRT product. Correspondingly, the CC values increased from GSMaP_NRT to GSMaP_Gauge_NRT, and the RMSE values showed an apparent downward trend. However, in terms of the contingency table statistics, the improvements were not obvious. This suggests that the calibration can effectively reduce the bias but is not good at improving the skill of detecting rainy events. Additionally, Figure 4 illustrates that the two near-real-time GSMaP products had different performances at four climate regimes, with better agreement from gauge observations over the humid region (Figure 4a,b) and an unsatisfactory performance over the arid region (Figure 4g,h). This was likely caused by the different retrieval skills of rainfall types over different climate regimes. The arid region, covered with desert and high mountains, was dominated by short-lived convective precipitation and orographic precipitation. However, the satellite-based precipitation retrieval had difficulty coping with these two precipitation conditions. Moreover, the light rainfall and winter snow in arid region further imposed another challenge to the satellite-based precipitation estimates. Consequently, the satellite precipitation products usually had an unsatisfactory performance over the arid region in China, and this is consistent with the result of Yong et al. [39] and Chen et al. [46]. Compared to GSMaP_NRT, the GSMaP_Gauge_NRT obviously improved the data accuracy, but it had low CC (0.36) and BIAS (35.08%) values over the arid region. Such results indicate that the current near-real-time gage calibration algorithm of GSMaP still have significant room for further improving the data quality over the arid region.

Theoretically, a value of satellite precipitation can be divided into four categories based on its ability to identify rain occurrences. A hit event means that both the satellite estimate and gauge reference detected rain, while miss precipitation suggests that a rain event was reported by gauge observation but not detected by satellite. In contrast, false precipitation means that precipitation was detected by satellite but not observed by gauge, and the rest part of precipitation means that that both satellite and gauge showed no rain. Based on the different precipitation events, we further computed the intensity distributions of daily precipitation amount to look into the error characteristics of GSMaP products (Figure 5). The rainfall intensity was binned with logarithmic scale across the range of 1–256 mm/day, and the daily averaged precipitation accumulation of each bin was calculated on the y-axis. The intensity distribution, which has different error components, can reveal detailed information on the error features at the event scale. As shown in Figure 5, the intensity distributions of total precipitation were generally similar to those of hit precipitation, suggesting that hit event was the dominate component of total precipitation. However, over the arid region, the false precipitation also accounted for a considerable proportion of the total precipitation. Considering that the false precipitation will amplify the total precipitation amount of satellite, the obvious overestimation of two near-real-time GSMaP products over the arid region was partly caused by the false rainy events. On the other hand, compared to GSMaP_NRT, the intensity distributions of GSMaP_Gauge_NRT were closer to the gauge observations (first column in Figure 5), which indicates that the GSMaP_Gauge_NRT has better performance than the GSMaP_NRT. Additionally, it can be seen that the calibration of GSMaP_Gauge_NRT mainly changed the intensity distributions at a moderate–high rain rate. For example, the GSMaP_NRT had more precipitation amounts than gauge observations over the semi-humid region (Figure 5e). After the calibration, the curve of GSMaP_Gauge_NRT descended and was more consist with the gauge. In the third column of Figure 5, it was found that the GSMaP_NRT and GSMaP_Gauge_NRT showed basically identical intensity distributions of miss precipitation over all four climate regions. This implies that the GSMaP_Gauge_NRT failed to correct precipitation events undetected by the satellites in the calibration process. Therefore, the future correction efforts of incorporating different precipitation component is recommended to improve the precision of satellite precipitation.
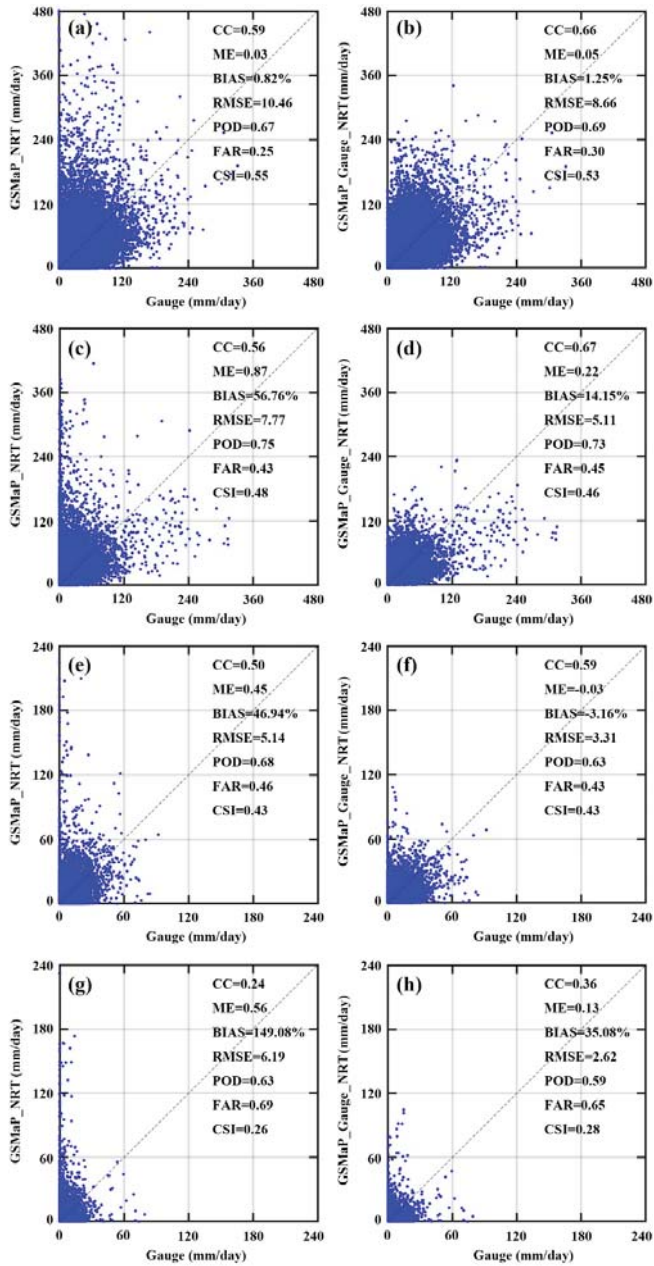
Figure 4. Scatterplots of the daily precipitation for GSMaP_NRT (left) and GSMaP_Gauge_NRT (right) versus gauge observations at selected grid boxes over four climate regions: (**a**,**b**) Humid region; (**c**,**d**) semi-humid region; (**e**,**f**) semi-arid region; (**g**,**h**) arid region.
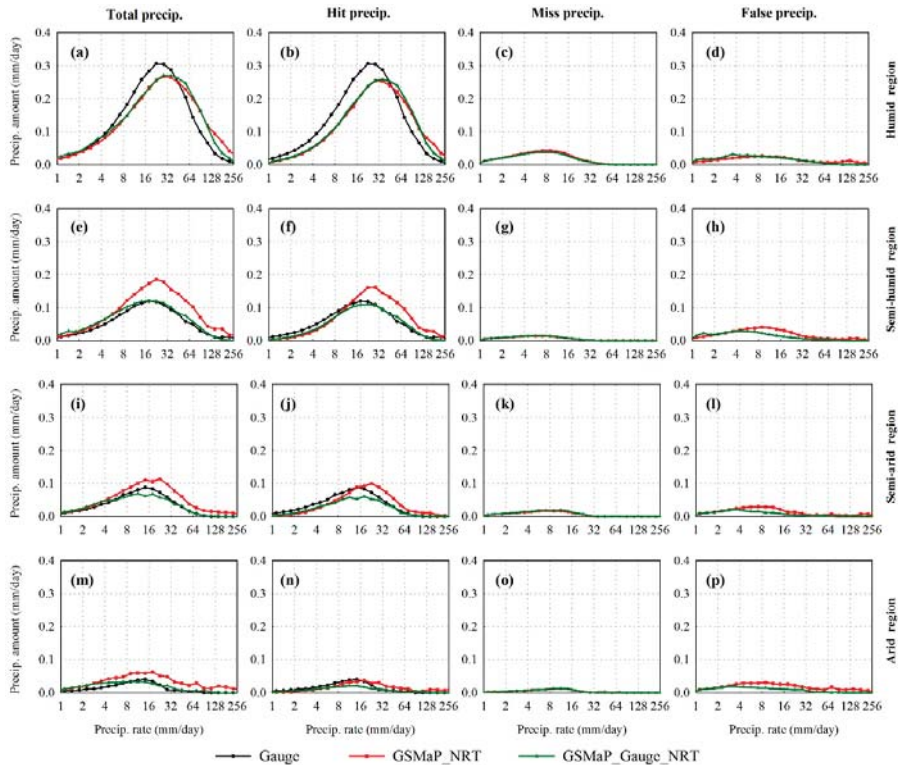
Figure 5. Daily intensity distribution of the total, hit, miss, and false precipitation over four climate regions: (**a–d**) Humid region; (**e–h**) semi-humid region; (**i–l**) semi-arid region; and (**m–p**) arid region. The total observed precipitation (black line) is also shown in the first two columns.

3.3. Global View of GSMaP_NRT and GSMaP_Gauge_NRT

In the previous section, we compared the GSMaP_NRT with GSMaP_Gauge_NRT satellite precipitation products over the Mainland China. The results showed that the GSMaP_Gauge_NRT effectively reduced the error in GSMaP_NRT and was more consistent with gauge observations. To obtain a much broader view of GSMaP_Gauge_NRT at other regions, we extended the comparison to the global scale. Operationally, it is usually difficult to validate the satellite precipitation products on the global scale using gauge observations, due to the fact that the ground gauge networks are sparse or nonexistent in many regions like oceans, deserts, and mountains. In order to provide a globally consistent evaluation, we chose the GSMaP_Gauge product as the reference data for global comparison. Figure 6 displays the spatial difference between the two near-real-time GSMaP products and the standard GSMaP_Gauge product. It can be seen that the GSMaP_NRT underestimated the precipitation in the southeastern China, which was consistent with the evaluation results when using CGDPA as a reference. The calibrated GSMaP_Gauge_NRT substantially decreased this underestimation and was closer to the GSMaP_Gauge. Globally, the GSMaP_NRT exhibited positive bias over most parts of the world. Most notably, along the intertropical convergence zone (ITPC) extending toward Central America, eastern United States, southern parts of South America, and West Africa, the GSMaP_NRT tended to overestimate precipitation. After the real-time calibration, these positive biases were effectively reduced. Besides, compared to the GSMaP_Gauge, the GSMaP_Gauge_NRT only had slight underestimations over most parts of world, except for the ITPC, with an overestimation (Figure 6b). Considering that the development of GSMaP_Gauge_NRT intended to make the GSMaP near-real-time

products as close to the GSMaP_Gauge product as possible, we believe that the parameterized gauge calibration effectively reduced the errors in GSMaP_NRT, and that the calibrated GSMaP_Gauge_NRT is a better product than GSMaP_NRT.
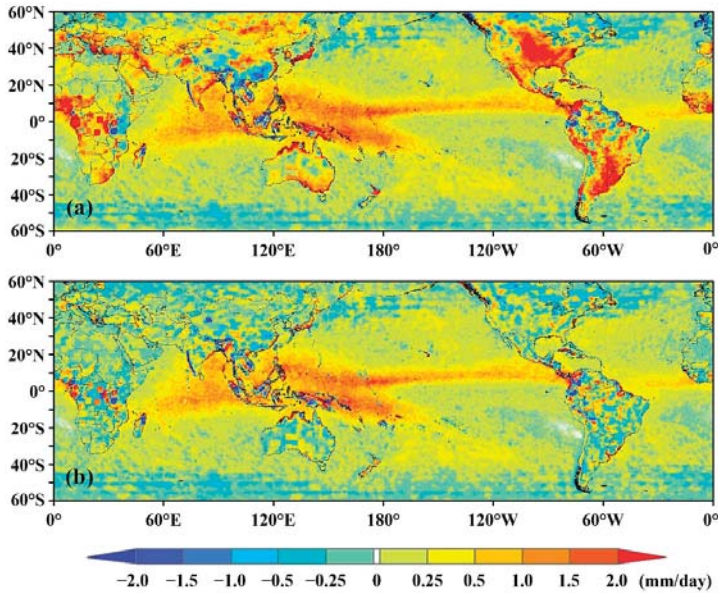

Figure 6. Global map of mean daily precipitation difference between (**a**) GSMaP_NRT and GSMaP_Gauge and (**b**) GSMaP_Gauge_NRT and GSMaP_Gauge.

4. Discussion

Remote sensing of precipitation provides an alternative source for precipitation data beyond ground observations. To date, numerous studies have reported the performance of satellite precipitation products in many areas around the world [20,53–58]. However, most of these evaluations focused on the post-real-time products that are usually available after a few days or months [59–61]. The near-real-time satellite precipitation products are more attractive for some application scenarios, owing to their timeliness. To the best of our knowledge, our study is a first evaluation of the gauge-adjusted near-real-time GSMaP precipitation estimate. The results in this study showed that the GSMaP_Gauge_NRT has a more reliable performance than its original GSMaP_NRT over China. This provides a positive feedback to the GSMaP algorithm team, as they expected. In the TRMM era, the CCA was proposed to improve the accuracy of TMPA-RT. However, this climatological calibration is not currently used in IMERG near-real-time products due to its poor performance at high latitudes and altitudes [9,31]. In the GPM era, the GSMaP and IMERG were the most popular high-resolution satellite precipitation products, and the GSMaP_NRT and IMERG-early were their near-real-time products with latency periods of 3 h and 4 h, respectively. Therefore, when the CCA is applied in the IMERG product, it is necessary to compare these two near-real-time products of GSMaP and IMERG, and this work can be continued in the future.

As a preliminary assessment of the near-real-time GSMaP precipitation estimate, we performed the validation at daily scale considering the availability of reference data. Previous studies found that the performances of satellite precipitation products are sensitive to the spatiotemporal scale [62,63]. Such a resampling in this study may lose the potential characters of GSMaP products at their original resolutions. However, our study mainly focused on the comparing of GSMaP_Gauge_NRT and GSMaP_NRT. Thus, this modification of scale should not change the conclusions of our study. In the

next work, we will investigate the performance of GSMaP_Gauge_NRT product in sub-daily or hourly time scales.

The evaluation showed that the calibrated GSMaP_Gauge_NRT has better performance than the original GSMaP_NRT. Thus, it is reasonable to choose the GSMaP_Gauge_NRT precipitation product for users with their requirement of near-real-time data. Of course, more researches are needed for knowing the application potentiality of GSMaP_Gauge_NRT. The performance of the IMERG near-real-time products have been investigated in capturing extreme precipitation events [64,65]. The results of these studies indicated that the IMERG has a promising potential for monitoring typhoon rainfall. Considering the accuracy improvement of GSMaP_Gauge_NRT, further application researches, like rainstorm monitoring and hydrological simulation, are encouraged. This will provide useful feedbacks and insights about the GSMaP_Gauge_NRT to decision-makers and the scientific community.

5. Conclusions

Recently, the GSMaP algorithm developers proposed a parameterized gauge calibration method to reduce the errors in GSMaP_NRT without jeopardizing its near-real-time availability. In this study, we compared and validated the calibrated GSMaP_Gauge_NRT product with the original GSMaP_NRT over the Mainland China, by using a high-quality ground gauge reference dataset.

Our analyses showed that the GSMaP_NRT product can well-capture spatial patterns of precipitation across the China, but it significantly overestimates the reference precipitation with BIAS of 15.84%. After bias adjustment, this overestimation was obviously reduced, with slight overestimation for GSMaP_Gauge_NRT (4.46%). Correspondingly, the value of CC rose from 0.58 for GSMaP_NRT to 0.67 for GSMaP_Gauge_NRT, and the RMSE was reduced from 9.11 mm to 7.07 mm. This indicates that the parameterized calibration strategy can effectively decrease the bias in the GSMaP_NRT, and that the calibrated GSMaP_Gauge_NRT has a better performance than the original GSMaP_NRT.

In terms of the contingency table statistics, we found that the improvements in the contingency table statistics were not obvious. This suggests that the calibration can effectively reduce the bias but is not good at improving the skill of detecting precipitation events. When we decomposed satellite precipitation into different rainy events, the results further validated that the correction scheme mainly occurred in the hit event and could hardly make up the rainfall missed by the satellites. Thus, we highlight that incorporation of precipitation components is of vital importance for future calibration work.

Finally, our evaluation was extended to the global scale to examine the performance of GSMaP_Gauge_NRT from a broader perspective. The global analysis showed that the bias in GSMaP_ NRT was generally alleviated after gauge calibration and the calibrated GSMaP_Gauge_NRT product was in good agreement with the GSMaP_Gauge product. Therefore, to summarize, all of the results in this study suggest that GSMaP_Gauge_NRT can effectively reduce the uncertainties in GSMaP_NRT after the calibration and that the GSMaP_Gauge_NRT is a more reliable near-real-time satellite precipitation product than the original GSMaP_NRT. As a preliminary assessment of GSMaP_Gauge_NRT product, we hope that this study provides useful information for algorithm developers and product users. Considering the diverse nature of the world's topography and climate characteristics, future studies are encouraged to evaluate and validate the performance of GSMaP_Gauge_NRT product in more regions using local density gauge networks.

the authors wish to extend their appreciation to the editors and four anonymous reviews for their thoughtful comments and insightful suggestions.

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

References

1. Michaelides, S.; Levizzani, V.; Anagnostou, E.; Bauer, P.; Kasparis, T.; Lane, J.E. Precipitation: Measurement, remote sensing, climatology and modeling. *Atmos. Res.* **2009**, *94*, 512–533. [CrossRef]

2. Kidd, C.; Huffman, G. Review global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]

3. 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]

4. 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]

5. Behrangi, A.; Khakbaz, B.; Jaw, T.C.; AghaKouchak, A.; Hsu, K.; Sorooshian, S. Hydrologic evaluation of satellite precipitation products over a mid-size basin. *J. Hydrol.* **2011**, *397*, 225–237. [CrossRef]

6. Massari, C.; Crow, W.; Brocca, L. An assessment of the performance of global rainfall estimates without ground-based observations. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 4347–4361. [CrossRef]

7. Sapiano, M.R.P.; Arkin, P.A. An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeorol.* **2009**, *10*, 149–166. [CrossRef]

8. Tapiador, F.J.; Turk, F.J.; Petersen, W.; Hou, A.Y.; García-Ortega, E.; Machado, L.A.T.; Angelis, C.F.; Salio, P.; Kidd, C.; Huffman, G.J.; et al. Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.* **2012**, *104*, 70–97. [CrossRef]

9. Yong, B.; Liu, D.; Gourley, J.J.; Tian, Y.; Huffman, G.J.; Ren, L.; Hong, Y. Global view of real-time TRMM multisatellite precipitation analysis: Implications for its successor global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 283–296. [CrossRef]

10. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.G.; Bowman, K.P.; Hong, Y.; Stocker, E.F.; Wolff, D.B. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]

11. 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]

12. Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Hirose, M.; Takayabu, Y.N.; Nakagawa, K.; Iwanami, K.; et al. Global precipitation map using satellite-borne microwave radiometers by the GSMaP Project: Production and validation. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2259–2275. [CrossRef]

13. Hsu, K.; Gao, X.; Sorooshian, S.; Gupta, H.V. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.* **1997**, *36*, 1176–1190. [CrossRef]

14. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A.; et al. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 150066. [CrossRef] [PubMed]

15. Ebert, E.E.; Janowiak, J.E.; Kidd, C. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 47–64. [CrossRef]

16. Tian, Y.; Peters-Lidard, C.D.; Adler, R.F.; Kubota, T.; Ushio, T. Evaluation of GSMaP precipitation estimates over the contiguous United States. *J. Hydrometeorpl.* **2010**, *11*, 566–574. [CrossRef]

17. Stampoulis, D.; Anagnostou, E.N. Evaluation of global satellite rainfall products over continental Europe. *J. Hydrometeorol.* **2012**, *13*, 588–603. [CrossRef]

18. Gebremichael, M.; Bitew, M.M.; Hirpa, F.A.; Tesfay, G.N. Accuracy of satellite rainfall estimates in the Blue Nile Basin: Lowland plain versus highland mountain. *Water Resour. Res.* **2014**, *50*, 8775–8790. [CrossRef]

19. Katsanos, D.; Retalis, A.; Michaelides, S. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. *Atmos. Res.* **2016**, *169*, 459–464. [CrossRef]

20. 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]

21. Nguyen, P.; Ombadi, M.; Sorooshian, S.; Hsu, K.; AghaKouchak, A.; Braithwaite, D.; Ashouri, H.; Thorstensen, A.R. The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 5801–5816. [CrossRef]

22. Khan, S.; Maggioni, V. Assessment of level-3 gridded Global Precipitation Mission (GPM) products over oceans. *Remote Sens.* **2019**, *11*, 255. [CrossRef]

23. Beck, H.E.; Pan, M.; Roy, T.; Weedon, G.P.; Pappenberger, F.; Van Dijk, A.I.J.M.; Huffman, G.J.; Adler, R.F.; Wood, E.F. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 207–224. [CrossRef]

24. 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]

25. Yong, B.; Hong, Y.; Ren, L.; Gourley, J.; Huffman, G.; Chen, X.; Wang, W.; Khan, S. Hydrologic evaluation of multisatellite precipitation analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. *Water Resour. Res.* **2010**, *46*. [CrossRef]

26. Chen, S.; Hong, Y.; Gourley, J.J.; Huffman, G.J.; Tian, Y.; Cao, Q.; Yong, B.; Kirstetter, P.E.; Hu, J.; Hardy, J.; et al. Evaluation of the successive V6 and V7 TRMM multisatellite precipitation analysis over the Continental United States. *Water Resour. Res.* **2013**, *49*, 8174–8186. [CrossRef]

27. Liu, Z. Comparison of precipitation estimates between Version 7 3-hourly TRMM Multi-Satellite Precipitation Analysis (TMPA) near-real-time and research products. *Atmos. Res.* **2015**, *153*, 119–133. [CrossRef]

28. Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Pai, D.S. Error characterization of TRMM Multisatellite Precipitation Analysis (TMPA-3B42) products over India for different seasons. *J. Hydrol.* **2015**, *529*, 1302–1312. [CrossRef]

29. Huffman, G.J. *The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG)*; NASA: Washington, DC, USA, 2018. Available online: https://pmm.nasa.gov/sites/default/files/document_files/TMPA-to-IMERG_transition_0.pdf (accessed on 6 September 2019).

30. Lu, D.; Yong, B. Evaluation and hydrological utility of the latest GPM IMERG V5 and GSMaP V7 precipitation products over the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 2022. [CrossRef]

31. Yong, B.; Ren, L.; Hong, Y.; Gourley, J.J.; Tian, Y.; Huffman, G.J.; Chen, X.; Wang, W.; Wen, Y. First evaluation of the climatological calibration algorithm in the real-time TMPA precipitation estimates over two basins at high and low latitudes. *Water Resour. Res.* **2013**, *49*, 2461–2472. [CrossRef]

32. 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]

33. Zhu, Z.; Yong, B.; Ke, L.; Wang, G.; Ren, L.; Chen, X. Tracing the error sources of global satellite mapping of precipitation for GPM (GPM-GSMaP) over the Tibetan Plateau, China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 2181–2191. [CrossRef]

34. Mega, T.; Ushio, T.; Takahiro, M.; Kubota, T.; Kachi, M.; Oki, R. Gauge-adjusted global satellite mapping of precipitation. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 1928–1935. [CrossRef]

35. Aonashi, K.; Awaka, J.; Hirose, M.; Kozu, T.; Kubota, T.; Liu, G.; Shige, S.; Kida, S.; Seto, S.; Takahashi, N.; et al. GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation. *J. Meteorolog. Soc. Jpn. Ser. II* **2009**, 119–136. [CrossRef]

36. Chen, Z.; Qin, Y.; Shen, Y.; Zhang, S. Evaluation of global satellite mapping of precipitation project daily precipitation estimates over the Chinese mainland. *Adv. Meteorol.* **2016**, *2016*, 9365294. [CrossRef]

37. Yong, B.; Wang, J.; Ren, L.; You, Y.; Xie, P.; Hong, Y. Evaluating four multisatellite precipitation estimates over the Diaoyu Islands during Typhoon seasons. *J. Hydrometeorol.* **2016**, *17*, 1623–1641. [CrossRef]

38. Satgé, F.; Hussain, Y.; Bonnet, M.P.; Hussain, B.M.; Martinez-Carvajal, H.; Akhter, G.; Uagoda, R. Benefits of the Successive GPM Based Satellite Precipitation Estimates IMERG–V03,–V04,–V05 and GSMaP–V06,–V07 Over Diverse Geomorphic and Meteorological Regions of Pakistan. *Remote Sens.* **2018**, *10*, 1373. [CrossRef]

39. Yong, B.; Chen, B.; Tian, Y.; Yu, Z.; Hong, Y. Error-component analysis of TRMM-based multi-satellite precipitation estimates over mainland China. *Remote Sens.* **2016**, *8*, 440. [CrossRef]

40. Xie, P.; Chen, M.; Yang, S.; Yatagai, A.; Hayasaka, T.; Fukushima, Y.; Liu, C. A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.* **2007**, *8*, 607–626. [CrossRef]

41. Shen, Y.; Xiong, A.; Wang, Y.; Xie, P. Performance of high-resolution satellite precipitation products over China. *J. Geophys. Res. Atmos.* **2010**, *115*. [CrossRef]

42. Shen, Y.; Xiong, A. Validation and comparison of a new gauge-based precipitation analysis over mainland China. *Int. J. Climatol.* **2016**, *36*, 252–265. [CrossRef]

43. Guo, H.; Chen, S.; Bao, A.; Behrangi, A.; Hong, Y.; Ndayisaba, F.; Hu, J.; Stepanian, P.M. Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China. *Atmos. Res.* **2016**, *176*, 121–133. [CrossRef]

44. Zhao, H.; Yang, S.; You, S.; Huang, Y.; Wang, Q.; Zhou, Q. Comprehensive evaluation of two successive V3 and V4 IMERG Final Run precipitation products over mainland China. *Remote Sens.* **2017**, *10*, 34. [CrossRef]

45. Sun, W.; Ma, J.; Yang, G.; Li, W. Statistical and Hydrological Evaluations of Multi-Satellite Precipitation Products over Fujiang River Basin in Humid Southeast China. *Remote Sens.* **2018**, *10*, 1898. [CrossRef]

46. Chen, H.; Yong, B.; Gourley, J.J.; Liu, J.; Ren, L.; Wang, W.; Hong, Y.; Zhang, J. Impact of the crucial geographic and climatic factors on the input source errors of GPM-based global satellite precipitation estimates. *J. Hydrol.* **2019**, *575*, 1–16. [CrossRef]

47. Tian, Y.; Peters-Lidard, C.D.; Eylander, J.B.; Joyce, R.J.; Huffman, G.J.; Adler, R.F.; Hsu, K.; Turk, F.J.; Garcia, M.; Zeng, J. Component analysis of errors in satellite-basd precipitation estimates. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

48. Li, Z.; Yang, D.; Hong, Y. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *J. Hydrol.* **2013**, *500*, 157–169. [CrossRef]

49. Kim, K.; Park, J.; Baik, J.; Choi, M. Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.* **2017**, *187*, 95–105. [CrossRef]

50. Xu, F.; Guo, B.; Ye, B.; Ye, Q.; Chen, H.; Ju, X.; Guo, J.; Wang, Z. Systematical Evaluation of GPM IMERG and TRMM 3B42V7 Precipitation Products in the Huang-Huai-Hai Plain, China. *Remote Sens.* **2019**, *11*, 697. [CrossRef]

51. You, Y.; Wang, N.Y.; Ferraro, R.; Rudlosky, S. Quantifying the snowfall detection performance of the GPM microwave imager channels over land. *J. Hydrometeor.* **2017**, *18*, 729–751. [CrossRef]

52. Ferraro, R.R.; Smith, E.A.; Berg, W.; Huffman, G.J. A screening methodology for passive microwave precipitation retrieval algorithms. *J. Atmos. Sci.* **1998**, *55*, 1583–1600. [CrossRef]

53. Dinku, T.; Ruiz, F.; Connor, S.J.; Ceccato, P. Validation and intercomparison of satellite rainfall estimates over Colombia. *J. Appl. Meteorol. Climatol.* **2009**, *49*, 1004–1014. [CrossRef]

54. 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. [CrossRef]

55. Duan, Z.; Liu, J.; Tuo, Y.; Chiogna, G.; Disse, M. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Sci. Total Environ.* **2016**, *573*, 1536–1553. [CrossRef] [PubMed]

56. Maggioni, V.; Massari, C. On the performance of satellite precipitation products in riverine flood modeling: A review. *J. Hydrol.* **2018**, *558*, 214–224. [CrossRef]

57. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K.-L. A review of global precipitation datasets: Data sources, estimation, and intercomparisons. *Rev. Geogr.* **2018**, *56*, 79–107. [CrossRef]

58. Prakash, S. Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *J. Hydrol.* **2019**, *571*, 50–59. [CrossRef]

59. Cattani, E.; Merino, A.; Levizzani, V. Evaluation of monthly satellite-derived precipitation products over East Africa. *J. Hydrometeorol.* **2016**, *17*, 2555–2573. [CrossRef]

60. Rozante, J.; Vila, D.; Barboza Chiquetto, J.; Fernandes, A.; Souza Alvim, D. Evaluation of TRMM/GPM blended daily products over Brazil. *Remote Sens.* **2018**, *10*, 882. [CrossRef]

61. Navarro, A.; García-Ortega, E.; Merino, A.; Sánchez, J.L.; Kummerow, C.; Tapiador, F.J. Assessment of IMERG Precipitation Estimates over Europe. *Remote Sens.* **2019**, *11*, 2470. [CrossRef]

62. Tan, J.; Petersen, W.A.; Kirstetter, P.E.; Tian, Y. Performance of IMERG as a function of spatiotemporal scale. *J. Hydrometeorol.* **2017**, *18*, 307–319. [CrossRef] [PubMed]

63. Omranian, E.; Sharif, H.O. Evaluation of the Global Precipitation Measurement (GPM) satellite rainfall products over the lower Colorado River basin, Texas. *J. Am. Water Resour. Assoc.* **2018**, *54*, 882–898. [CrossRef]

64. Su, X.; Shum, C.K.; Luo, Z. Evaluating imerg v04 final run for monitoring three heavy rain events over mainland China in 2016. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 444–448. [CrossRef]

65. Huang, C.; Hu, J.; Chen, S.; Zhang, A.; Liang, Z.; Tong, X.; Xiao, L.; Min, C.; Zhang, Z. How Well Can IMERG Products Capture Typhoon Extreme Precipitation Events over Southern China? *Remote Sens.* **2019**, *11*, 70. [CrossRef]

Article

Ground Validation of GPM IMERG Precipitation Products over Iran

Fatemeh Fadia Maghsood [1,2], Hossein Hashemi [1], Seyyed Hasan Hosseini [1,3,*]

1   Center for Middle Eastern Studies & Department of Water Resources Engineering, Lund University,
    SE-221 00 Lund, Sweden; fatemeh.maghsood@cme.lu.se (F.F.M.); hossein.hashemi@tvrl.lth.se (H.H.);
    ronny.berndtsson@tvrl.lth.se (R.B.)
2   Department of Watershed Management and Engineering, College of Natural Resources, Tarbiat Modares
    University, Tehran 1411713116, Iran
3   Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz 5166616471, Iran
*   Correspondence: seyyed_hasan.hosseini@cme.lu.se

Abstract: Accurate estimation of precipitation is crucial for fundamental input to various hydrometeorological applications. Ground-based precipitation data suffer limitations associated with spatial resolution and coverage; hence, satellite precipitation products can be used to complement traditional rain gauge systems. However, the satellite precipitation data need to be validated before extensive use in the applications. Hence, we conducted a thorough validation of the Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals (IMERG) product for all of Iran. The study focused on investigating the performance of daily and monthly GPM IMERG (early, late, final, and monthly) products by comparing them with ground-based precipitation data at synoptic stations throughout the country (2014–2017). The spatial and temporal performance of the GPM IMERG was evaluated using eight statistical criteria considering the rainfall index at the country level. The rainfall detection ability index (POD) showed that the best IMERG product's performance is for the spring season while the false alarm ratio (FAR) index indicated the inferior performance of the IMERG products for the summer season. The performance of the products generally increased from IMERG-Early to –Final according to the relative bias (rBIAS) results while, based on the quantile-quantile (Q-Q) plots, the IMERG-Final could not be suggested for the applications relying on extreme rainfall estimates compared to IMERG-Early and -Late. The results in this paper improve the understanding of IMERG product's performance and open a door to future studies regarding hydrometeorological applications of these products in Iran.

Keywords: GPM IMERG; Iran; satellite precipitation; spatiotemporal analysis; statistical distribution; validation

1. Introduction

Precipitation plays a crucial role in the Earth's hydrological cycle and is a fundamental input to a wide range of hydrological, meteorological, and climate model applications [1,2]. Thus, accurate estimation of the precipitation amount and pattern is vital for improved prediction of water-related processes as well as reducing uncertainties for effective water resource management practices [3,4]. To obtain precipitation amounts, ground-based measurements, i.e., rain gauges and weather radars, are considered a reliable source mainly at the local scale. At the regional and global scale, however, there are limitations for using ground-based measurements, particularly in most developing countries [5]. Radar networks are often available where there is a coverage by rain gauges. However, radars are subject to different errors and uncertainties, such as ground clutter, anomalous propagation, signal attenuation, beam blockage, and bright band contamination [6].

Rain gauges are limited in describing the spatial distribution of precipitation depending on the arrangement and density of the rain gauge network [7,8]. In order to spatially characterize precipitation, gauge measurements are transformed to a gridded precipitation dataset. This is carried out through interpolation of rain gauge measurements, using spatial interpolation and geo-statistical methods [9]. These may be prone to missing values, wind effects, insufficient numbers of rain gauges, and a sparse network, especially in less accessible mountainous and oceanic areas [4].

In view of the above, the spatial limitations, resolution, and coverage of ground-based measurements highlight the importance of satellite-based precipitation estimates at both the regional and global scale. Satellite-based precipitation estimates are also subject to uncertainties through cloud top reflectance, thermal radiance, infrequent satellite overpasses, and retrieval algorithm related to the nature of indirect measurement [10]. Therefore, a thorough validation of satellite precipitation data in any given area is necessary to achieve insight regarding is accuracy as well as identifying sources of errors to improve algorithms and satellite sensor development. Further, accuracy assessment taking into account the pros and cons of satellite precipitation estimates is imperative before using data in hydrological modeling in any given region [11,12]. Such findings help in selecting a supportive product for a special application under different circumstances [1].

Given the success of the Tropical Rainfall Measuring Mission (TRMM), the National Aeronautical and Space Administration (NASA) and Japan Aerospace Exploratory Agency (JAXA) launched a new generation Global Precipitation Measurement (GPM) mission in early 2014 to replace the TRMM mission [13]. The GPM mission is expected to compensate the limitations of TRMM precipitation products by providing higher resolution, larger spatial coverage, and more accurate global precipitation estimates [14]. The GPM precipitation algorithm, Integrated Multi-satellite Retrievals for GPM (IMERG), is based upon the experiences from the TRMM algorithm. As the spatiotemporal resolution and coverage of GPM have been extended beyond the TRMM resolution and coverage, the performance of the GPM IMERG products needs to be evaluated and validated globally.

Several studies have compared the GPM IMERG and TRMM products with ground-based measurements, i.e., rain gauge and weather radar [4,6,10,15–19], considering their hydrological applications [14,20–22]. Also, different GPM IMERG products regarding temporal resolutions have been evaluated considering various climatic and topographic conditions using various statistical measures across the world [5,23–30]. Although most of these studies confirmed the improvement of the IMERG products relative to those of the TRMM Multi-satellite Precipitation Analysis (TMPA), a more comprehensive investigation is still essential to better understand the IMERG performance in various regions of the world taking into consideration different products' versions and temporal resolution. Countries in the Middle East suffer from acute hydrometeorological data shortage, both in terms of quality and quantity [15], and Iran is not an exception. Rain gauges are sparse and unevenly distributed throughout the country, particularly in remote areas of the center and eastern areas. Delays in data processing and publishing for public access and scientific use and an absence of data sharing in many trans-boundary basins constitute a main shortcoming for ground-based precipitation data in the country [5]. To our knowledge, there are very few investigations of the IMERG products´ performance over Iran on a basin scale [5,10]. There are no comprehensive studies that investigate the performance of the IMERG product at the country level.

According to the above, the newly available IMERG products have not been thoroughly explored for Iran as a whole. The country covers different climatic, geographic, and topographic features, with respect to temporal and spatial particularities and different satellite products´ versions. This study aimed to provide a better understanding of the IMERG product's performance over the country and open a door to future studies regarding hydrological and hydrometeorological applications of these products at both the local and regional scale. Accordingly, we performed a comprehensive evaluation of the performance of IMERG products considering three time-latencies, IMERG-Early, IMERG-Late, and IMERG-Final, and two temporal resolutions, daily and monthly, based on eight criteria indices. We examined these criteria in view of spatial and temporal patterns related to features,

such as elevation, slope, latitude, and longitude, over the entire Iran. Also, the statistical distributions of the precipitation products were compared to that of ground measurements for different seasons.

2. Materials and Methods

2.1. Study Area

Iran covers an area of about 1.648 million km^2, located between 44–64° East and 25–40° North in the eastern part of the Middle East [31]. The country is bordered by Azerbaijan, Armenia, Turkmenistan, and the Caspian Sea to the north; Afghanistan and Pakistan to the east; Oman Sea and Persian Gulf to the south; and Turkey and Iraq to the west (Figure 1a). The temporal and seasonal distribution of precipitation is governed by the interaction of the tropical air mass from the Red Sea, the Mediterranean low pressure, and the Siberian and western high pressures [31].
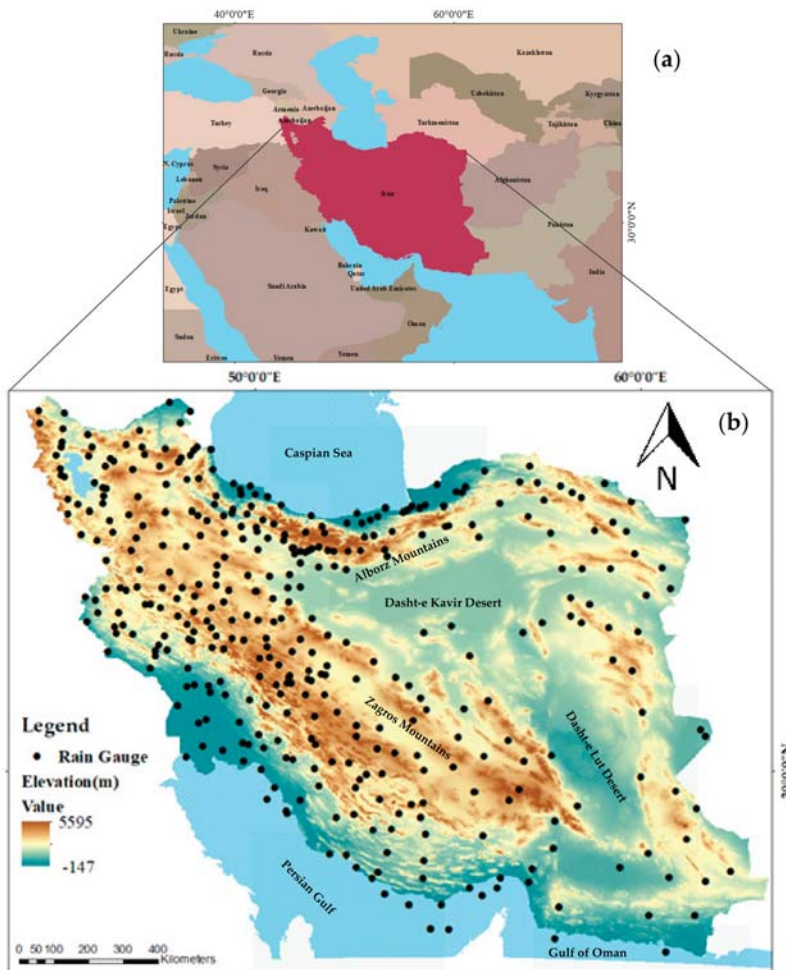

Figure 1. (**a**) Map of Iran and (**b**) distribution of rain gauges with altitude.

Half of Iran's land surface is mountainous, a quarter is covered by fertile and productive plains, and the final quarter by salty arid desert [32]. Zagros and Alborz are the two main mountainous areas,

which are located along the western and northern borders of the country, respectively. These mountain ranges prevent much of the available humidity, mostly initiating from the Mediterranean and Caspian Sea in the West and North of the country, from reaching the interior parts, so singularly govern the spatial patterns of precipitation across Iran [31,33] (Figure 1b). As a result, the central parts of the country and the southwestern areas of the Caspian Sea receive an average of 100 and 1800 mm year-1, respectively [31].

Iran encircles a variety of climates, which range from extremely to subtropical dry in the central to eastern parts of the country, wet to extremely wet in the coastal plains of the Caspian Sea, relatively wet in some western areas, and arid to semiarid zones in the rest of the country [34]. The Mediterranean Synoptic System is the dominant weather system in the country, where the rainy season is from October to April, with a peak in December. The driest month is July.

2.2. Datasets

2.2.1. Satellite-Based Precipitation (GPM IMERG)

The GPM mission was launched on 27 February 2014 by NASA and JAXA as an international joint project for frequent measurements of near-global precipitation. Like the antecedent mission TRMM, the GPM is a constellation of multiple satellites, which comprise the GPM Core Observatory carrying combined passive/active sensors, and microwave measurements by partner satellites. The Core Observatory information acts as a reference standard to combine other satellites' microwave precipitation measurements that orbit within the same constellation. The GPM Core Observatory technically consists of two major sensors: GPM Microwave Imager (GMI) and Dual-frequency Precipitation Radar (DPR). The GMI sensor measures precipitation characteristics, such as intensity, type, and size, while DPR observes the internal structure of storms within and under the clouds and measures precipitation in 3D throughout the atmospheric column and provides an insight into the structure of rain drops [7].

The success of the GPM mission depends upon algorithm development that combines both GMI and DPR observations and the partner satellites' measurements. GPM data are available online at three various processing levels, including geo-located and calibrated (Level 1), geophysical data product derived from Level 1 data (Level 2), and a composite of Level 2 data products (Level 3) [35]. The GPM Level-3, namely IMERG, is derived from multiple satellites and available as three different types of daily products consisting of early run (IMERG-Early), late run (IMERG-Late), and final run (IMERG-Final), which are published with 4-h, 12-h, and 2.5-month latencies, respectively, and a monthly product (IMERG-Monthly). These products are accessible at $0.1° \times 0.1°$ spatial resolution for regions between 60° S and 60° N globally and at several temporal resolutions, i.e., half-hourly, 3-hourly, daily, and monthly [15]. More detailed information and an algorithm description can be found in [35]. While IMERG-Early provides a quick estimate with only a 3-h latency period, IMERG-Late successively presents better estimates as data from more partner satellites is merged with a 12-h latency period. IMERG-Final estimates are, however, presented after bias adjustment with monthly rain-gauge measurements with a three-month latency. The IMERG-Early data can be used for potential flood or landslide warnings while the IMERG-Late and Final data can be used in agricultural forecasting, drought monitoring, and hydrological modeling [24].

It is noted that the IMERG-Final run algorithm provides two different types of precipitation estimates, including precipitationCal (with rain gauge-adjusted processes) and precipitationUnCal (without rain gauge-adjusted processes). Based on the literature, precipitationCal has indicated better performance relative to ground-based measurements [25]. Therefore, this algorithm together with the IMERG-Early, -Late, and -Monthly products from April 2014 until December 2017 were used in this study.

2.2.2. Ground-Based Precipitation

The ground-based precipitation data were obtained from the I.R. IRAN Meteorological Organization (IRIMO). IRIMO is responsible for recording hourly and daily precipitation and quality control is carried out before releasing recorded data to the user's community across the country [33]. This dataset is considered as the most reliable precipitation data source in Iran [31].

Although these data are freely available for research purposes, the data are not available online and should be obtained in person. In total, 403 meteorological synoptic stations were under operation across the country, of which only 370 stations had at least one year of daily records and were selected for this study. Perfectly, 368 and 349 of the selected stations had at least two and three years of daily records for the studied period, respectively, between April 2014 and December 2017. While GPM IMERG data were historically complete (without missing data), the selection of this period is attributed to the availability of both satellite and rain gauge precipitation datasets. Figure 1b shows the distribution of the synoptic stations throughout the country.

2.3. Evaluation Processes and Indices

2.3.1. Data Preparation

The GPM IMERG products that cover the entire country from 1 April 1 2014 to 31 December 2017 at daily and monthly time scales and $0.1° \times 0.1°$ (about 11 km on the equator) resolution were acquired. The RT-H5 file format, in which RT refers to real-time and H5 denotes an HDF5 file, was converted to the ASCII format and all days were stacked by a written algorithm in R programming language. The precipitation was converted to mm day^{-1} for all IMERG products for the purpose of consistency with ground-based data. As for the point (rain-gauges) to pixel (IMERG products) comparison, each IMERG pixel corresponding to single rain gauge locations was extracted. The extracted IMERG data and the corresponding rain gauge data for daily and monthly time scales were then used in the analyses.

2.3.2. Error Analysis of IMERG Products

To assess the performance of satellite-based precipitation products against measured precipitation by rain gauges, widely applied evaluation indices, including the mean absolute error (MAE), Pearson correlation coefficient (CC) [14], and relative bias (rBIAS) [36], were used. In this regard, MAE (Equation (1)) indicates the error distribution and mean magnitude of errors without considering direction. MAE has the same unit as the precipitation data (i.e., mm day^{-1}). These criteria were calculated for each grid covering the attributed rain gauges. Daily and monthly products were separately analyzed, mainly due to the smooth nature of monthly data and superior performance relative to daily products. The CC (Equation (2)) shows the agreement between the precipitation estimated by the satellite and rain gauge measurements. CC is a dimensionless number, which varies between −1 and 1, with CC equal to zero when there is no correlation. The rBIAS (Equation (3)) represents the size and direction of the difference between the two datasets. Positive and negative rBIAS is an overall overestimation and underestimation of the satellite relative rain gauge measurements, respectively [7]. MAE and rBIAS close to 0 and CC close to 1 display the best performance of the IMERG products relative to the rain gauge measurements in this study:

$$\text{MAE} = \frac{\sum_{i=1}^{n} |S_i - O_i|}{n},\tag{1}$$

$$\text{CC} = \frac{\sum_{i=1}^{n} \left(S_i -, \overline{S}\right)\left(O_i, -, \overline{O}\right)}{\sqrt{\sum_{i=1}^{n}\left(S_i, -, \overline{S}\right) \sum_{i=1}^{n}\left(O_i, -, \overline{O}\right)}},\tag{2}$$

$$\text{rBIAS} = \frac{\sum_{i=1}^{n}(S_i - O_i)}{\sum_{i=1}^{n} O_i} \ or \ \frac{\overline{S} - \overline{O}}{\overline{O}}, \tag{3}$$

where O_i and S_i are the observed rain gauge and satellite-based precipitation data, respectively, $\overline{O}$ and $\overline{S}$ are the rainfall averages for pixel i associated to the rain gauge, and n is the total number of satellite-gauge data pairs, which are being compared.

In addition, to investigate how often a significant over/under-estimation by the satellite takes place regardless of the overall magnitude and direction of the errors, we introduced two new indices, named over and under, based on introducing a preliminary index of equal, which stands for an insignificant error. This corresponds to an error smaller than 10% as compared to measurements. However, errors smaller than 0.25 mm day^{-1} were considered insignificant as well. The over, under, and equal indices are presented as percentages.

Further, to quantify the precipitation detection ability of the satellite-based precipitation estimates against the ground-based observations, two indices, including the probability of detection (POD) and false alarm ratio (FAR) were calculated (Equations (4) and (5)). The POD expresses the ratio of the correct precipitation detection of the satellite and FAR measures the proportion of no-rain events that are recorded as rain by the satellite. The closest values to 100% and 0% display the best satellite performance for POD and FAR, respectively [14]. The following equations define the POD and FAR:

$$\text{POD} = \frac{n_{11}}{n_{11} + n_{10}} \times 100, \tag{4}$$

$$\text{FAR} = \frac{n_{01}}{n_{11} + n_{01}} \times 100, \tag{5}$$

where n_{11} is the number of rainfall events that are observed by the rain gauge and detected by the satellite, n_{10} is the number of rainfall events that are observed by the rain gauge but not detected by the satellite, and n_{01} is the number of rainfall events that are detected by the satellite but not observed by the rain gauge.

2.3.3. Analysis of Statistical Distribution

The statistical distribution of rainfall data is an important component of hydrological and hydrometeorological studies, such as in intensity–duration–frequency (IDF) relationships and design storms. The quantile-quantile (Q-Q) plot is a graphical tool for determining whether the two datasets, i.e., ground- and satellite-based, have similar distributional shapes. The technique is conducted by plotting quantiles (or percentiles) of the two datasets versus one another and comparing the plot with a 45° reference line. Accordingly, the Q-Q plot is a scatter plot, with the points falling approximately along the reference line standing for a common distribution for the two datasets. On the contrary, the greater the departure from the reference line, the greater the evidence for refusing this assumption. It is worth noting that the quantiles of a dataset are the points below which a certain proportion of the data lies. For example, in a classic standard normal probability distribution with a mean of 0, the 0.5 quantile (or 50th percentile), 0 means that half the data are not exceeding 0.

There are also analytical methods, such as the chi-square and Kolmogorov–Smirnov 2-sample tests, that are used for assessing if two sets of quantiles follow the same distribution. However, the Q-Q plot is favorable as it provides more insight into the nature of the difference between two datasets than analytical methods. Although the Q-Q plot is only a visual check rather than an air-tight proof, it helps to observe if the assumption is plausible and, otherwise, which data points at which quantile cause the violation of the assumption.

The Q-Q plot can easily show the under/overestimation of a dataset, i.e., satellite, compared to the rain gauge, between percentiles of the datasets. Additionally, many distributional aspects, including shifts in location, shifts in scale, change in symmetry, tail behavior, and the presence of outliers, can be discerned. The behavior of the tail of the Q-Q plot can be important for extreme hydrology studies.

The tail refers to data points associated to statistically rare incidents, such as values above the 95th or 99th percentile of the datasets.

2.3.4. Geospatial and Temporal Analysis

The difference between the IMERG products and rain gauge measurements may depend on geospatial conditions, time of the year, and the rainfall nature at the different geospatial locations. To investigate this dependency, the evaluation indices presented in Section 2.3.2 were calculated for different categories of geospatial factors, such as the elevation, slope, latitude, and longitude of satellite grids corresponding to rain gauges locations. The evaluation was carried out for additional factors, including the average annual rainfall, mean dry period (intervals between rainfall events), and rainfall index (average annual rainfall/mean dry period) for stations representative of the frequency and amount of rainfall at a given location. For simplicity, a location with a lower or higher rainfall index is, alternatively, mentioned as a dryer or wetter location, respectively.

We used box plots to illustrate the variation of the indices for 10 different categories of rain gauges. These categories were based on selecting 11 range limits of a given geospatial factor so that an equal number of rain gauges was placed into each category (37 rain gauges out of the selected 370 synoptic stations). Therefore, the 1st and the 11th limits were the minimum and the maximum values of the geospatial factor among the selected locations of the study.

Further, the spatiotemporal evaluations were separately reported for each month and season using combined maps. That is, plotting the spatial variation of the indices, e.g., POD, and, at the same time, the factors, e.g., topography, for a given month or season in a single map.

3. Results and Discussion

### 3.1. General Comparison of Precipitation Datasets

Monthly, seasonal, and annual average precipitation from daily IMERG-Early, -Late, -Final, and -Monthly products together with the corresponding ground observations are shown in Figure 2, Figure 3, and Figure 4. We defined the seasons as winter (Dec-Jan-Feb), spring (Mar-Apr-May), summer (Jun-Jul-Aug), and fall (Sep-Oct-Nov). Figure 2 shows the monthly averages of daily products, including IMERG-Early, -Late, -Final, and -Monthly, compared to rain gauge-measured precipitation during the study period. It can be seen that the difference between IMERG-Early, -Late products, and observations is maximum in April and May (Figure 2). The rain gauge corrected IMERG-Final and -Monthly products showed good agreement with observations during these two months. The correlation between IMERG-Early and -Late products and ground observations in Figure 2 is 0.93 and 0.99 for IMERG-Final and -Monthly products, respectively. In other words, the IMERG-Early and -Late daily products are generally less accurate in estimating the monthly average rainfall amounts as compared to the IMERG-Final. Further, the comparison indicates that IMERG-Final and -Monthly are highly correlated during all months and consistent with the rain gauge measurements.

According to Figure 3, the IMERG-Early and -Late slightly underestimated the measured precipitation in winter (by −7% and −12%, respectively) and overestimated the measured precipitation in the spring (by 33% and 25%, respectively). In contrary, the IMERG-Final and -Monthly overestimated the rain gauge measurements in winter (by 6% and 12%, respectively) but adequately matched measured precipitation in the spring (overestimation by ~5%). Considering summer and fall, all daily products showed high performance for the entire country while the IMERG-Monthly product performed well for spring, summer, and fall with overestimation in the winter. Overall, IMERG-Final and -Monthly products showed high performance in estimating seasonal precipitation relative to measurements.

Average annual precipitation (Figure 4) showed no major discrepancy between the IMERG-Late and -Final (overestimation by only 2% and 1%, respectively) products relative to rain gauge measurements, but the IMERG-Early and -Monthly products overestimated by 7% and 5%, respectively, relative to the measurements.
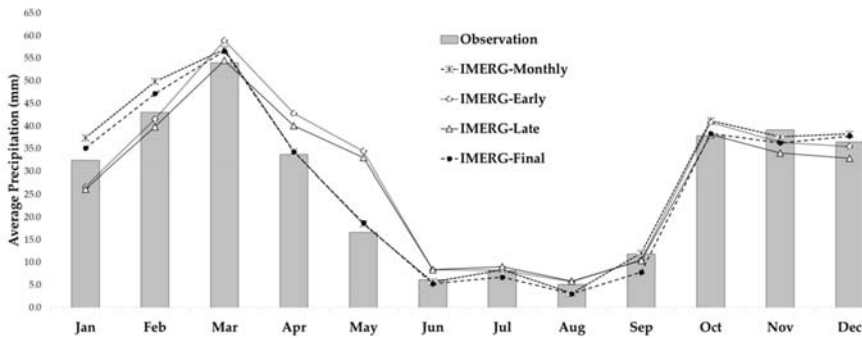
Figure 2. Monthly averages of IMERG-Early, -Late, and -Final based on the daily time step and -Monthly products compared to ground-based precipitation measurements (Observation).
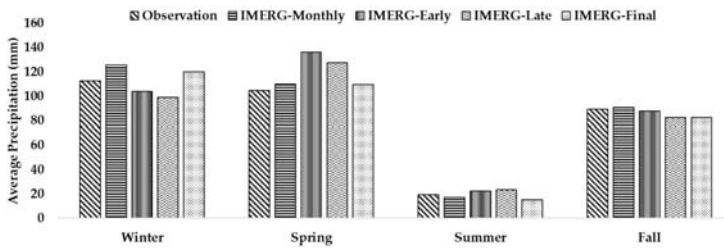
Figure 3. Seasonal averages of monthly and daily precipitation products (IMERG-Early, -Late, and -Final), compared to ground-based precipitation measurements (Observation).
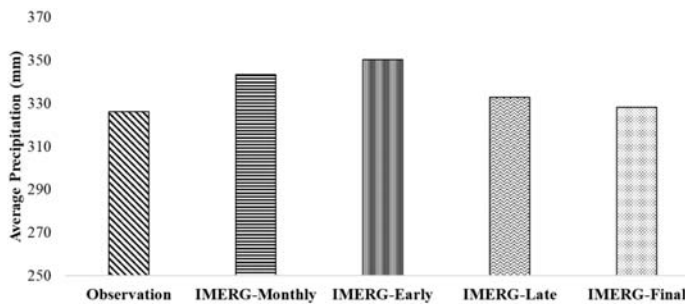
Figure 4. Annual averages of monthly and daily precipitation products (IMERG-Early, -Late, and -Final), and ground-based precipitation measurements (Observation).

3.2. Error Analyses of IMERG Products

A comparison of corresponding spatiotemporal rainfall data in the two datasets, i.e., satellite and rain gauge from the same location and time, was carried out using the evaluation indices introduced in Section 2.3.2. Table 1 presents these indices, including CC, MAE, rBIAS, POD, FAR, Equal, Over, and Under for the daily (IMERG-Early, -Late, and –Final) and monthly (IMERG-Final) time scale over the country. The indices were calculated based on the overlaid IMERG pixels and rain gauge data coordinates during the 2014–2017 period. For the comparison purposes, the daily products were accumulated for each month and presented in the monthly time scale, so all products can be compared in a monthly time scale supported with different values for the criteria indices monthly time scale presented in Table 1. It is noted that the algorithms used for producing the IMERG-Monthly product are different from those of the IMERG-Final, although both are gauge-corrected products.

Table 1. Error analyses for daily and monthly IMERG products vs. rain gauge measurements in both monthly and daily time scales.

Criteria	IMERG-Monthly	IMERG-Early		IMERG-Late		IMERG-Final	
	Monthly	Monthly	Daily	Monthly	Daily	Monthly	Daily
CC	0.68	0.49	0.41	0.50	0.41	0.65	0.47
MAE (mm day^{-1})	0.44	0.62	1.14	0.60	1.10	0.43	1.03
rBIAS	0.05	0.04	0.09	−0.01	0.03	−0.04	0.00
POD (%)	99.8	97.9	67.6	98.3	68.0	98.3	67.9
FAR (%)	22.7	14.6	60.1	15.5	59.9	15.5	59.6
Over (%)	34.9	38.7	46.2	35.6	44.9	31.1	46.1
Equal (%)	47.1	36.3	12.0	38.2	12.7	48.3	13.2
Under (%)	18.0	25.0	41.8	26.3	42.3	20.6	40.7

Generally, in the monthly time scale, the IMERG-Monthly product showed the best performance in comparison to rain gauge measurements, in relation to the daily IMERG products, especially for IMERG-Early and IMERG-Late. However, the IMERG-Final products indicated, generally, comparable results and even, to some degree, outperformed the IMERG-Monthly product according to the MAE, rBIAS, FAR, Over, and Equal criteria. The CC between monthly rainfall from gauges and satellite products was the highest, 0.68 for the IMERG-Monthly, and the MAE was the lowest, yielding 0.43 mm day^{-1} for the IMERG-Final. In the case of relative bias between rain gauges and satellite products, the results were excellent, yielding less than ±5% except for IMERG-Early in the daily time scale. The bias calculation showed that the IMERG-Early products slightly overestimated the rain gauge measurements at the country level, but a substantial improvement regarding rBIAS can be seen from IMERG-Early, 9% to Final, 0% in daily time scale (Table 1).

Generally, in the case of IMERG daily products, in the daily time scale, criteria indices for the three products did not significantly differ from each other, but the performance of the daily IMERG products relative to the rain gauge measurements improved from IMERG-Early to -Final. The IMERG-Final showed the highest correlation with rain gauge measurements, with a CC value of 0.47, followed by the IMERG-Late and -Early both yielding 0.41. The MAE and rBIAS decreased from the IMERG-Early to -Final products, showing a better performance of the IMERG-Final product relative to –Early and –Late in comparison with the rain gauge measurements in the daily time scale (Table 1).

The results were acceptable for the satellite precipitation detection ability in all daily IMERG products: POD = 0.68% in the daily time scale and POD = 98%–100% in the monthly time scale. The POD did not change from the IMERG-Early to -Final daily products, meaning that the satellite-gauge-adjusted algorithm, used in the IMERG-Final product, did not help for the detection improvement of the IMERG-Early and –Late products. As mentioned in Section 2.2.2, the GPM Core Observatory's temporal resolution is three hours, which results in no calibration of the precipitation observed by the constellation satellites within this gap for a given area. However, this problem is resolved by accumulating the daily into monthly precipitation (POD = 1). Similar results were achieved for FAR calculations. The IMERG daily products showed low performance by having a false detection value of 0.6 while the monthly FAR was 0.23, indicating an acceptable rain detection ability at the monthly time scale (Table 1).

At the daily time scale, in principal, no significant improvement in rainfall detection ability regarding POD and FAR indices were observed. Further, no substantial differences between all three daily products were observed, as the monthly product showed a better performance in the estimation of precipitation considering MAE and rBIAS, and precipitation detection, POD, and FAR.

The over index did not change much from IMERG-Early to -Final and the under index decreased slightly from IMERG-Early (42%) to -Final (41%) products. For the daily products, equal was much smaller than both over and under, and over was 2% to 6% higher than under. However, for the monthly product, equal was about 47%, and over was 17% higher than under (Table 1). The more smoothed nature of the monthly data variation, as compared to daily, is a probable reason for the higher equal

obtained for the monthly product. Also, the larger upper values as compared to under for all products is in line with the overall overestimation shown by the positive rBIAS. However, the cause of a larger difference between upper and under for the monthly product, compared to the daily ones, can be explained under some circumstances. For instance, suppose a month with only one (or few) rainy day is overestimated by the satellite. While, at a daily time scale, this will be counted as only one (a few) day of overestimation in the calculation of over, at the monthly time scale, this would be counted as one month as compared to the length of monthly data, which is almost 30 times shorter than the daily ones, therefore, resulting in higher over.

3.3. Evaluation of Statistical Distribution of IMERG Products

Figure 5 displays the Q-Q plots for each IMERG daily product and different seasons, including winter (December-February), spring (March-May), summer (June-August), and fall (September-November). It should be noted that these plots are not showing data from the tails of the distributions (close to the 100th percentile). For clarity, the Q-Q plots contain rainfall data at least until the 95.5th percentile. As a result, only rainfall lower than 37 mm day^{-1} (typical high rainfall amount for Iran) was taken into consideration, allowing a comparison between all Q-Q plots. In addition, the exact location of some typical quantiles is depicted by an arrow and a percentile value to better visualize the data distribution.

As an example, Figure 5a shows that the most obvious departures for the IMERG-Early product in comparison to rain gauge observations in winter start above the 90th percentile of data (i.e., rainfall >3 mm day^{-1}) in the form of an underestimation that gradually increased until the 97.5th percentile (~14 mm day^{-1}). This underestimation remained rather constant until the 99.5th percentile (~36 mm day^{-1}) by an amount of about 2.5 to 3 mm day^{-1} while there was a slight overestimation for the more frequent rainfall events (lower than the 90th percentile), with rainfall amounts less than 1 mm day^{-1}. Figure 5b shows that the IMERG-Late, to a lesser extent, overestimated low rainfall, but larger underestimations are seen for higher rainfall compared to the IMERG-Early product. However, as shown in Figure 5c, the higher accuracy of the IMERG-Final product leads to a significant reduction of departures from the 45-degree line that is observed for the IMERG-Early and -Late products in the form of underestimation. This result shows an acceptable fit for rainfall located approximately between the 95th and 99th percentile (~6–24 mm day^{-1}) while the overestimation extends up to the 95th percentile (i.e., rainfall <7 mm day^{-1}). That could be the reason for the overestimated average winter rainfall for the IMERG-Final product observed in Figure 3.

In Figure 5d, the scatter of points is linearly located to, and slightly above, the 45-degree line for rainfall less than the 99.5th percentile. This can be seen as a shift in the location of the statistical distribution of the IMERG-Early products in comparison to the distribution of rain gauge data. In other words, the IMERG-Early overestimated daily rainfall for values less than 29 mm in spring by an amount of about 1 to 1.5 mm. As seen in Figure 5e, the overestimation was reduced for the IMERG-Late product such that the points between the 97.5th and the 99.5th percentiles nearly overlaid the 45-degree line. On the other hand, as seen in Figure 5f, while the IMERG-Final product resulted in a better fit for the more frequent daily rainfalls in spring (i.e., less than 11 mm), major departures for higher values in the form of underestimation appear close to the 97.5th percentile. Therefore, a better fit for frequent rainfall by IMERG-Final suggested by Figure 5f seems to be the reason for the best estimation of average spring rainfall depicted in Figure 3. This product cannot be used as a firm reference dataset for extreme rainfall studies. Instead, a corrected IMERG-Early or -Late product, by removing the shift value for every data point, will be a better choice for the evaluation of extreme rainfall events in spring, which recently caused severe flood events in the country [37].
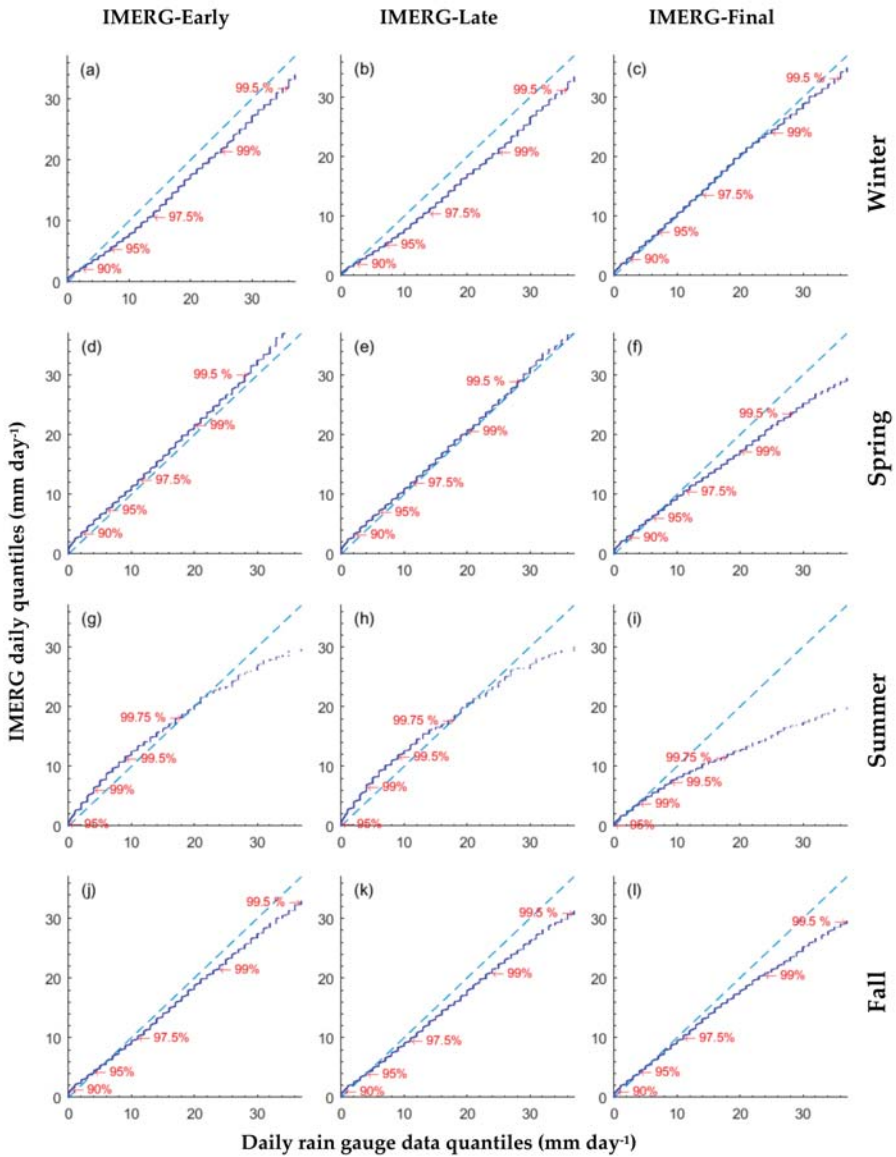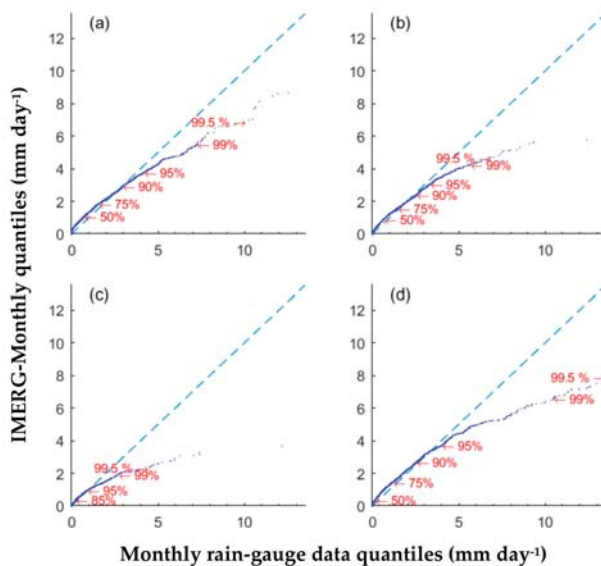
Figure 5. Q-Q plots of daily rainfall (IMERG and rain gauge) during 2014–2017: (**a**) IMERG-Early for Winter, (**b**) IMERG-Late for Winter, (**c**) IMERG-Final for Winter, (**d**) IMERG-Early for Spring, (**e**) IMERG-Late for Spring, (**f**) IMERG-Final for Spring, (**g**) IMERG-Early for Summer, (**h**) IMERG-Late for Summer, (**i**) IMERG-Final for Summer, (**j**) IMERG-Early for Fall, (**k**) IMERG-Late for Fall, and (**l**) IMERG-Final for Fall. The arrows and percentile values in red depict the location of quantiles.

The worst distributional fit between IMERG daily products and rain gauge observations is observed for summer (Figure 5g–i). Among all daily products, the best fit with measurements in summer for the data lower than the 99th percentile (~2 mm day^{-1}) is observed for the IMERG-Final (Figure 5i). However, the tail behavior for summer (>10 mm day^{-1}) for the IMERG-Final indicates a

large discrepancy relative to the other products (above 99.5th percentile). Finally, for the fall season (Figure 5j–l), all products indicate an overestimation of daily rainfall less than the 95th percentile and an underestimation of daily rainfall above the 95th percentile. However, the magnitude of overestimation is slightly reduced for IMERG-Late and IMERG-Final relative to IMERG-Early. While the IMERG-Final is the best product for fall daily rainfall lower than the 97.5th percentile, the underestimation deteriorates for extreme rainfall (obviously, above the 99th percentile) (Figure 5l). As a common result, while the bias-adjusted GPM IMERG products, which is the case for IMERG-Final datasets, resulted in better match with the gauge measurements for more frequent rainfall events (lower amounts of rainfall), uncorrected datasets of IMERG-Early and –Late products were shown to be more trusted related to the extreme events (heavy rainfall and flooding), especially in spring, summer, and fall seasons, as the bias-corrected data from the IMERG-Final product deteriorated underestimations observed for extreme rainfalls.

Figure 6a–d show Q-Q plots for the IMERG-Monthly product relative to rain gauge measurements for different seasons. The increasing underestimation from the 45-degree line for data above the 95th percentile for all seasons indicates that the IMERG-Monthly product might not be the best choice for the study of extreme monthly rainfall. It is noteworthy that the amount of extreme rainfall varies by season. For example, the average rainfall for the 99.5th percentile based on rain gauge measurements was about 4.3 and 13.5 mm day^{-1} for summer and fall, respectively. Based on the IMERG-Monthly product, however, the equivalent values for these months were 2 and 7.8 mm day^{-1}. This implies an underestimation as big as 69 mm for summer and 171 mm for the fall months (considering a 30-day month) that can obviously bias extreme monthly rainfall study based on the IMERG-Monthly product across the country. For more frequent rainfall (below the 95th percentile), especially in spring (Figure 6b) and summer (Figure 6c), the distribution of monthly data from the IMERG-Monthly product adequately fits the measurements, with an exception of overestimation for data lower than the 75th percentile, especially for winter (Figure 6a), fall (Figure 6d), and spring (Figure 6b).

Figure 6. Q-Q plots of average monthly rainfall (IMERG and rain gauge data) during the 2014–2017 period for (**a**) winter, (**b**) spring, (**c**) summer, and (**d**) fall seasons. The arrows and percentile values in red depict the location of quantiles.

3.4. Temporal Performance of IMERG Products

To explore the temporal characteristics of the calculated criteria for daily and monthly IMERG products, a monthly-based comparison was conducted using radar charts. It is noted that due to the similarity of the results for IMERG-Final and -Monthly, the monthly radar charts are separately displayed in the Supplementary Materials.

Figure 7a–h show a comparison of the eight criteria indices, including CC, rBIAS, MAE, FAR, POD, Under, Equal, and Over for each month considering IMERG-Early, -Late, and -Final products. The results show that, in general, there is a major temporal discrepancy in the criteria values obtained for different months. In most of the calculated criteria indices, the IMERG-Final shows noticeable improvements relative to IMERG-Early and -Late as compared to the rain gauge measurements.
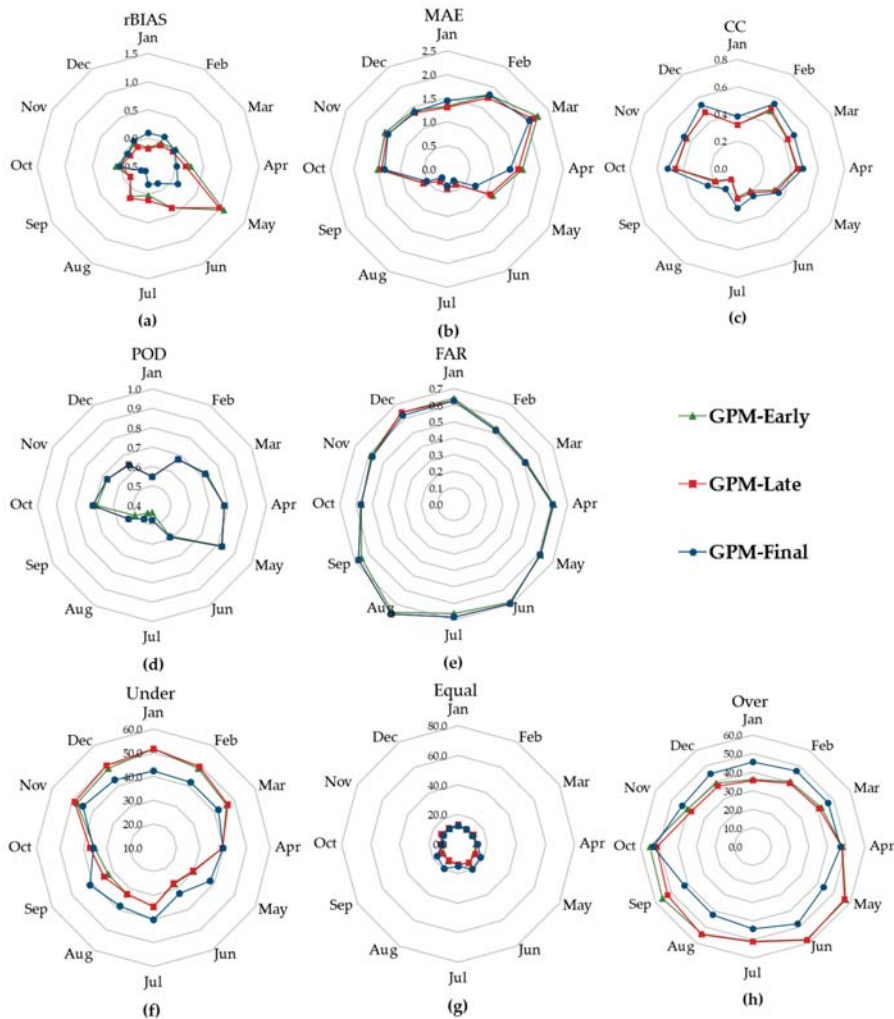

Figure 7. Radar charts for criteria indices for IMERG daily products at the monthly time scale.

In the case of rBIAS (Figure 7a), it can be seen that the largest overestimation is in May for IMERG-Early (1.07) and IMERG-Late (0.98) while the overestimation reduced to 0.13 for the

IMERG-Final product. It seems that the applied corrected algorithm considerably reduced the bias for the Final product relative to Early and Late products. All three products showed largest MAE in March and the lowest MAE in July (Figure 7b). The lowest CC is observed for the summer months, June, July, and August, for all products (Figure 7c) while the highest CC is obtained for February (0.49), December (0.48), and October (0.46), which is consistent with the rBIAS results. The weak performance of the IMERG products with regard to the detection ability of the satellite relative to gauges in the summer months is noticeable in the POD and FAR indices (Figure 7d,e). The detectability of the satellite is the highest from February to May, with POD between 0.68 and 0.82. The lowest FAR is obtained for the same months, between 0.51 and 0.62 (Figure 7d,e). The under, equal, and over values show that from IMERG-Early to -Final, the percentage of the under index increased from May to August for which the over index decreased (Figure 7f–h). On the other hand, the under index decreased from November to March while the over index increased.

3.5. Geospatial Performance of IMERG Products

Several studies have revealed the relationship between physical parameters, i.e., elevation, slope, latitude, and longitude, temperature etc., and satellite observation error [36]. Some of the errors are associated with the satellite sensor technology and applied algorithm while others are related to physical parameters on the ground. In the following section, we discuss the most prominent results obtained from the investigations of the relationships between rainfall index factors, introduced in Section 2.3.4, with multiple statistical criteria (i.e., rBIAS, CC, POD, etc.) that can introduce errors to the satellite data accuracy. The resulting boxplots for other geospatial factors can be found in the Supplementary Materials.

Relation to Location-Specific Rainfall

The charts presented in Figure 8 display the variation of location-specific criteria indices of IMERG daily products for different categories of stations based on the rainfall index factor, which, among all factors, showed a significant relationship with criteria indices. For these boxplots, limits of the bins (categories of rainfall index) are presented in the horizontal axis. For example, a hypothetical station with an average annual rainfall of 200 mm and a mean dry period of 20 days (i.e., rainfall index of $200/20 = 10$) is located in the third bin. Accordingly, an equal number of stations fall inside each bin. The vertical axis shows the variation of a specific index, e.g., CC. Figure 9 represents the location map of the stations in each category using different colors. As shown in this map, rain gauges with the highest rainfall index are located in the northern regions, mostly adjacent to the Caspian Sea coastline, as well as in the western regions. However, by moving from north and west to the central, eastern and southern regions, the rainfall index generally decreases. This spatial pattern is mainly controlled by the effect of two major mountain ranges in Iran (Alborz along the northern and Zagros along the western borders).

In Figure 8a–h, there are three boxplots for each bin in blue, green, and red corresponding to the IMERG-Early, -Late, and -Final product, respectively. According to the CC chart (Figure 8a), for instance, the IMERG-Final showed a higher correlation with rain gauge measurements in comparison to both IMERG-Early and –Late. These two displayed rather identical variation in each bin. Also, for the first and the last bins, there is a tendency to a lower correlation in comparison to other categories although the variation of CC in each bin is rather high. In general, the CC between the rain gauge data and satellite products in the daily time scale varies between 0 and 0.9 for different stations in the country.

As seen in Figure 8b–d, generally, by increasing the rainfall index (i.e., for stations in wetter locations), a lower frequency of overestimations (Figure 8b) and higher frequency of underestimations (Figure 8d) appear for all products. However, no significant change is observed for the equal index (i.e., negligible difference) at different bins (Figure 8c).

By comparing all three charts (Figure 8b–d), it appears that the frequency of overestimations (over) for stations in the first bin (i.e., the driest locations) decreases accompanied by increase in the

frequency of underestimations (under) and, to some degree, the frequency of negligible differences (equal) of the IMERG-Final product compared to the other daily products. For stations in the last bin, the condition for over and under was reversed while the frequency of negligible differences (equal) increased again for the IMERG-Final product. It can be concluded that the correction process by the IMERG-final product results in different changes in the frequency of over- and underestimations for different locations in Iran while it provides an overall decrease in the error for a majority of locations. The latter statement is confirmed by looking at Figure 8f,h as the alterations in the IMERG-Final boxplots in each bin relative to the boxplots of two other products are mostly close to a reduced MAE, as well as a reduced magnitude (absolute value) of rBIAS.
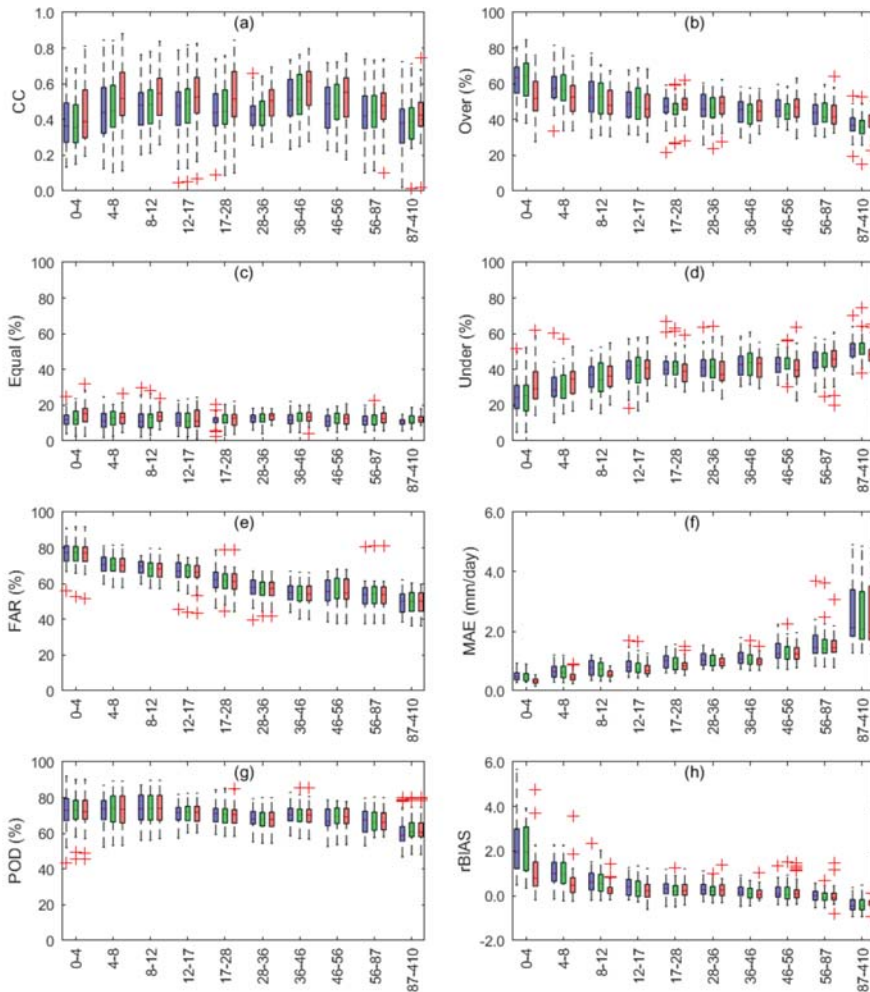
Figure 8. Box plots of the criteria indices for 10 rainfall index bins in blue, green, and red corresponding to IMERG-Early, IMERG-Late, and IMERG-Final daily products, respectively. The horizontal line in the boxes, and the upper and lower bounds of the boxes are the 50th, 75th, and 25th percentiles, respectively. The red plus symbols denote the outlier data and the whiskers (dashed black lines) extend to the most extreme data not considered as outliers.
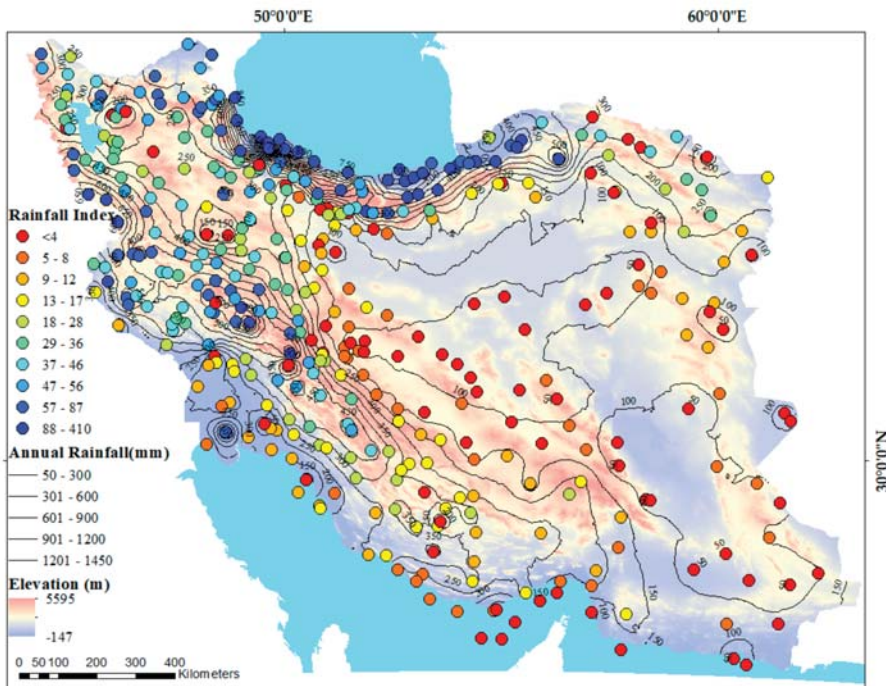
Figure 9. Spatial distribution of rain gauges colored differently for different categories of the rainfall index on the map of digital elevation and average annual rainfall contour lines.

MAE is the average magnitude of individual errors, so smaller MAE is favorable. However, it can result in a misleading interpretation. For example, at a dry location with zero rainfall for more than 90% of the entire length of the dataset, MAE will not reflect if there are a few major individual errors related to the extreme events. On the other hand, rBIAS calculates the accumulated individual errors (overall bias) relative to the accumulated observed rainfall during the period of comparison. Thus, it represents both overall under and overestimations (according to the negative or positive sign) and a comparable bias for different locations. As a result, the use of MAE together with rBIAS is essential. While a small magnitude of both MAE and rBIAS indicates a high performance of the satellite products, a combination of a large rBIAS with a low MAE for a location can be interpreted as a typical low individual error. This situation is more likely to appear for dryer locations with a higher frequency of smaller rainfall amounts. Also, a low rBIAS needs to be considered in the case of a large MAE value. Figure 10 illustrates these statements using Q-Q plots for a few locations selected from the different categories of the rainfall index.

In theory, a low rBIAS means that the total amount of rainfall observed at a location is accurately estimated by the satellite during the period of comparison. In other words, the sum of the positive individual errors is almost equal to the sum of the absolute values of the negative individual errors, regardless of the magnitude of the individual errors. On the other hand, CC is an accuracy criterion showing the degree of linear correlation between two datasets, thus it is not aimed to be an error index. To be able to discuss how these indices may result in contradictory situations, Figure 11 shows Q-Q plots for five other locations, with different combinations of CC, MAE, and rBIAS. It can be discerned that at locations no. (1) and no. (2), deviations from the 45-degree line are smaller, compared to the other locations, so the satellite product showed a better performance relative to location no. (3) and an even higher performance relative to locations no. (4) and no. (5) in presenting the actual daily rainfall

distribution. The correlation values for locations no. (1) and no. (2) are substantially different. On the other hand, the performance of the IMERG-Early product for the locations no. (3) and no. (5) seems to be completely different from each other while they both showed a high correlation (0.79) and a low MAE (~0.7 mm day^{-1}). Therefore, rBIAS could play a more discriminating role than a misleading CC or MAE, in the comparison of satellite-gauge datasets for their statistical distribution.
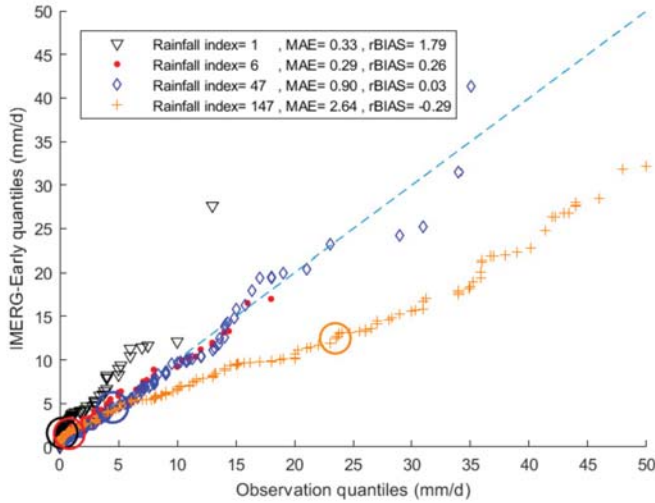
Figure 10. Q-Q plots for comparing different combinations of rBIAS and MAE at different locations with variable rainfall indices where lower values of MAE and lower absolute values of rBIAS are favorable (the large circles' center shows the position of the 95th percentile).
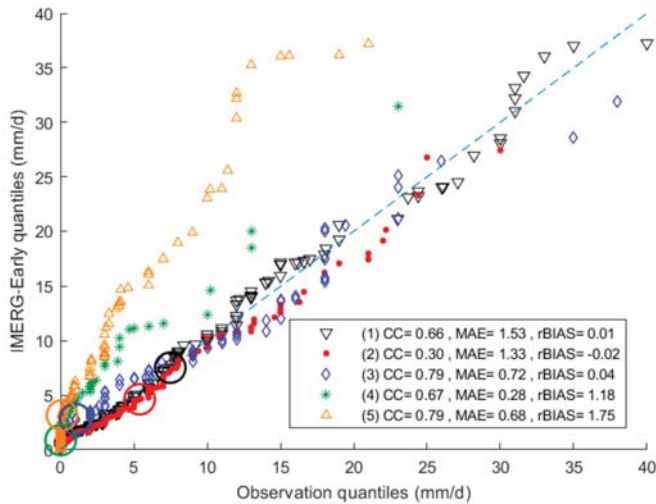

Figure 11. Q-Q plots comparing different combinations of CC, rBIAS, and MAE at five different locations where very low values of rBIAS show better performance for different combinations of CC and rBIAS (large circles' center shows the position of the 95th percentile).

To be able to evaluate the detection ability of the satellite products, the calculation of FAR and POD criteria is necessary (Figure 8e,g). As seen in Figure 8e, generally, FAR values were reduced with an increase of the rainfall index. The FAR values were higher for the majority of locations in the driest category (the first bin) as compared to wetter locations. For example, a median value of FAR at about 75% for the first category means that 75% of the rainy days detected by the satellite were not observed by the rain gauge. Also, the overall minimum value of FAR of around 35%, mostly, for the locations in the wetter locations indicated that, at least, 35% of the rainfall events detected by the satellite were not recorded by the rain gauges located within the corresponding satellite grids across the country. Regardless of errors due to the interruption of the rain gauge measurements or false detection by the satellite sensor, which are both possible, the increase of FAR for the dryer bins (Figure 8e) suggests that local rainfall events are more likely to appear at dryer locations. For such conditions, a rainfall event that partially affects a grid may not necessarily be observed by a rain gauge located in a dry part of the grid. Conversely, the chance for this condition is reduced for the wetter location, where uniform rainfall over a vast area is common.

According to Figure 8g, POD for different locations in Iran varied between 45% and 95% and more frequently between 60% and 80%. The higher PODs were more frequent at dryer locations and less frequent at wetter locations. POD indicates the chance for the satellite to detect a rainfall event, which is observed by a rain gauge within the satellite grid. According to this definition, the POD is not related to the spatial variability of rainfall in a grid. Instead, it indicates the sensor's inability to detect rainfall due to the temporal variability and the satellite visiting time. The variation of FAR and POD was almost the same for different daily products of IMERG, hence the applied correction in the IMERG-Final product did not account for the detection ability of the sensors. It appears that the IMERG corrections to the final product are mostly targeting the bias in the satellite observation. There is, however, some consistency in the results of different criteria. For example, the highest frequency of underestimations (Figure 8d) and negative rBIAS values (Figure 8h) for the wettest locations (locations in the last bin) can share common reasons related to the detection problems as the lowest values of POD were observed for a majority of the locations located in the last bin (Figure 8g).

The boxplots comparing the IMERG-Monthly products' performance for different categories of location (based on rainfall index) showed similar trends in the variation of the criteria indices for different categories of the rainfall index factor (Figure 12). The correlation for most of the locations was above 0.7 (Figure 12a). The over generally indicated a decreasing trend with the rainfall index (Figure 12b) similar to what was observed for the daily products while the frequencies of negligible differences (equal) between the IMERG-Monthly product and the rain gauge measurements showed a decreasing trend by the increase in the rainfall index (Figure 12c). The frequencies of the underestimations also showed an increase by the rainfall index (Figure 12d). The FAR decreased sharply compared to the daily FAR (Figure 12e). However, there are still considerable FAR values (i.e., above 40%) for the first three bins in Figure 12e (drier locations), which is related to local rainfall events in summer months (when there are only a few rainy days). This implies that the rainfall events are not uniformly distributed over a given satellite grid so the rain gauge located in the grid cannot, in some months, record any rainfall, but the satellite sensor does. MAE in Figure 12f shows an increase in the rainfall index, with a gentle slope compared to the increasing trend observed for the daily products (Figure 8f), which can be due to the smoother nature of the monthly data compared to the more erratic daily rainfall data. The POD for monthly data was close to 100% for almost all locations (Figure 12g), because there is a high chance that both the rain gauge and satellite recorded at least a rainy day in a given month. For the rBIAS, there is overestimation for almost all locations, with a rainfall index between 0 and 87, while the satellite underestimated monthly rainfall for a majority of the location in the 10th bin (the wettest category of locations) in Figure 12h.
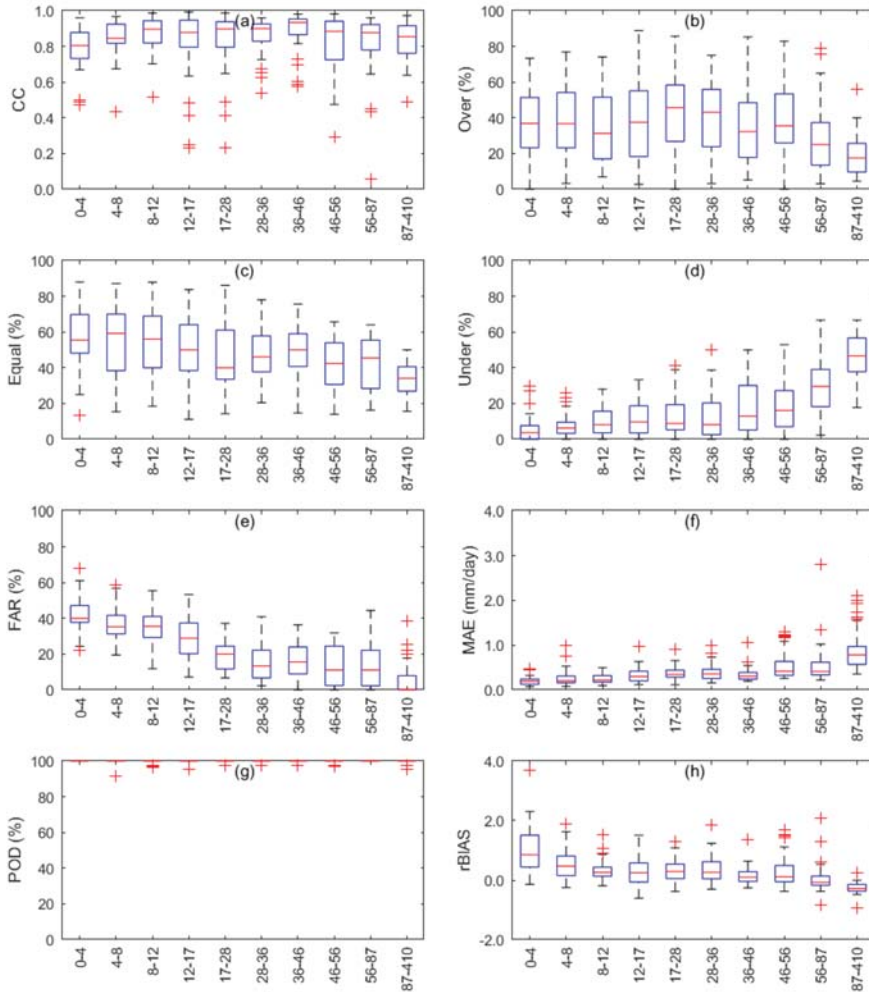
Figure 12. Box plots of the criteria indices for the 10 rainfall index bins for the IMERG-Monthly product. The horizontal line in the boxes, and the upper and lower bounds of the boxes are the 50th, 75th, and 25th percentiles, respectively. The red plus symbols denote the outlier data and the whiskers (dashed black lines) extend to the most extreme rainfall data not considered outliers.

3.6. Spatiotemporal Distribution of Rainfall Detection Ability

According to the results discussed earlier, the error indices varied in different parts of Iran. Since the variation of the error indices, such as rBIAS and MAE, can partly be explained by the rainfall detection ability of the satellite sensors. In this section, the rainfall detection ability of the satellite precipitation products are further assessed through spatial maps of POD and FAR. Figure 13 illustrates the spatial pattern of POD and FAR for four different seasons over the country. It is noted that the three IMERG daily products exhibit a similar spatial pattern of criteria relative to the rain gauge measurements with a slightly higher accuracy and lower bias for the IMERG-Final. Hence, we only discuss the criteria indices obtained from the analyses of the IMERG-Final and rain gauge

measurements in this section. These criteria indices were mapped using Kriging method in Arc GIS 10.4.1 environment.

According to the POD spatial map over the country (Figure 13), the satellite performance regarding precipitation detection shows an acceptable performance in most parts of the country in spring, followed by fall and winter (POD > 0.5). On the contrary, in the summer season, the southwestern parts, which show the best POD in winter, spring, and fall, indicate a low performance of the satellite in precipitation detection (POD < 0.4). It should be mentioned that the southern part of the country receives the end of the monsoon during summer time, for which precipitation is characterized by high intensity short-term rainfall [38]. Since the GPM constellation satellites revisit a given spot approximately every three hours, there is a high possibility that some of these short-term events are not observed by the satellite but by the rain gauges. As discussed in [1,39], higher POD is typically observed in dryer areas, i.e., central deserts (Figure S8), and the lower POD are typical for coastal areas, which is consistent with our findings, i.e., Persian Gulf and Caspian Sea coastal regions for summer (Figure 13).
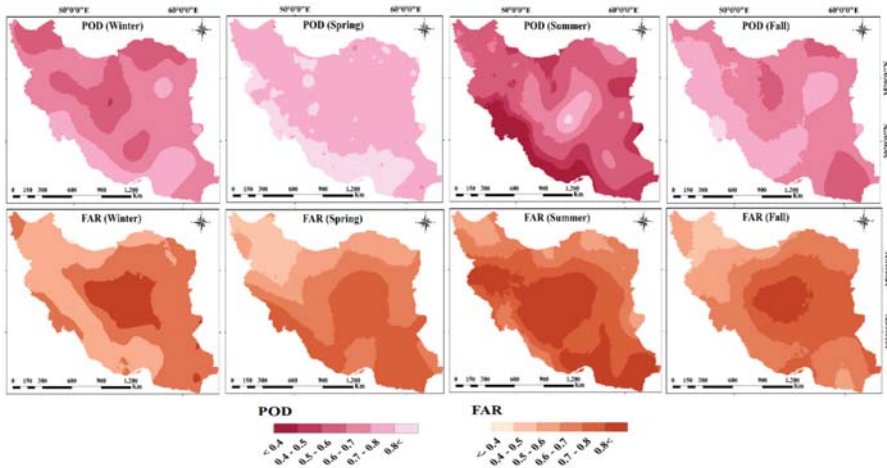

Figure 13. Spatial distribution of POD and FAR for different seasons during the 2014–2017 period.

According to the FAR spatial map of the country, the northwestern followed by the western portion of the country shows lower FAR in winter, spring, and fall. Similar to the POD spatial map, higher FAR is obtained in summer for these regions. In all seasons, the central part of the country shows the highest FAR confirmed by a sparse rain gauge network (see Figure 1b) in the central deserts, i.e., Kavir and Lut deserts that cover the dry and extremely dry zones (Figure S8). As mentioned before, the FAR implies the ratio between the number of rain events that are observed by the satellite but not recorded by the rain gauges.

4. Conclusions

In this study, the performance of IMERG GPM products was evaluated at a daily (Early, Late, and Final) and monthly temporal resolution using a high-quality rain gauge network over Iran during 2014 to 2017. The study is one of the first IMERG GPM product assessments at a country level taking into account temporal and geospatial properties. In this regard, the study used eight criteria indices, including CC, MAE, rBIAS, POD, FAR, Under, Over, and Equal. Additional analyses were carried out based on these indices taking into account temporal and geospatial features.

The general performance of IMERG products relative to the rain gauge measurements indicated a major improvement in the IMERG accuracy from IMERG-Early to -Final products. However, the two

indices of precipitation detection ability, POD and FAR, presented no major changes from Early to Final, which means that the correction algorithms do not account for the temporal correction of the satellite estimates. To evaluate the statistical distribution of rain gauge measurements versus satellite products, the Q-Q plots conclude that the IMERG-Final is not the best choice in extreme rainfall studies, but the IMERG-Early or Late can be used instead. Besides, the temporal performance of IMERG products, as displayed in the radar charts, showed a reduction of rBIAS from IMERG-Early to –Final.

Regarding POD, the best and worst performances were found in the spring and summer seasons, respectively. The FAR radar charts indicated an inferior performance of satellite products during the summer season.

The investigation of the relationship between various physical factors and location-specific factors of rainfall (rainfall index) with the eight mentioned criteria indices showed that CC varied for different rainfall indices. It appears that lower CC values were achieved both in the wettest and the driest locations. Further, by the increase of the rainfall index (from dryer to wetter locations), a lower and higher frequency of overestimation and underestimation, respectively, was observed for all IMERG products. Also, higher values of FAR were detected for the majority of the driest category of locations relative to wetter locations. Higher values of POD were found to be more frequent at dryer locations. As the POD investigated the spatial variability of rainfall within a particular grid, the results confirmed the superior detection ability of satellite sensors relative to gauge measurements (point measurement).

In general, the performance of satellite products increased from IMERG-Early to -Final products at the country level; however, these products need to be validated at the local scale and implemented in various hydrological models for verification. Higher values of FAR in the central part of the country, which is subjected to a sparse rain gauge network, require more caution when the IMERG data products are to be implemented in local-scale studies. This study provides an insight regarding the performance of the GPM IMERG products over all of Iran and can be used as a reference for further examination of the IMERG products in various hydrometeorological and hydrological applications.

References

1. Katiraie-Boroujerdy, P.S.; Nasrollahi, N.; Hsu, K.L.; Sorooshian, S. Evaluation of satellite-based precipitation estimation over Iran. *J. Arid Environ.* **2013**, *97*, 205–219. [CrossRef]
2. Mahmoud, M.T.; Al-Zahrani, M.A.; Sharif, H.O. Assessment of global precipitation measurement satellite products over Saudi Arabia. *J. Hydrol.* **2018**, *559*, 1–12. [CrossRef]

3. Sun, W.; Sun, Y.; Li, X.; Wang, T.; Wang, Y.; Qiu, Q.; Deng, Z. Evaluation and Correction of GPM IMERG Precipitation Products over the Capital Circle in Northeast China at Multiple Spatiotemporal Scales. *Adv. Meteorol.* **2018**, *2018*, 4714173. [CrossRef]

4. Tan, M.; Duan, Z. Assessment of GPM and TRMM Precipitation Products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]

5. Sharifi, E.; Steinacker, R.; Saghafian, B. Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sens.* **2016**, 135. [CrossRef]

6. Islam, T.; Rico-Ramirez, M.A.; Han, D.; Srivastava, P.K.; Ishak, A.M. Performance evaluation of the TRMM precipitation estimation using ground-based radars from the GPM validation network. *J. Atmos. Sol. Terr. Phys.* **2012**, *77*, 194–208. [CrossRef]

7. Kim, K.; Park, J.; Baik, J.; Choi, M. Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.* **2017**, *187*, 95–105. [CrossRef]

8. De Coning, E. Optimizing satellite-based precipitation estimation for nowcasting of rainfall and flash flood events over the South African domain. *Remote Sens.* **2013**, *5*, 5702–5724. [CrossRef]

9. Zhao, H.; Yang, B.; Yang, S.; Huang, Y.; Dong, G.; Bai, J.; Wang, Z. Systematical estimation of GPM-based global satellite mapping of precipitation products over China. *Atmos. Res.* **2018**, *201*, 206–217. [CrossRef]

10. Khodadoust Siuki, S.; Saghafian, B.; Moazami, S. Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products. *Int. J. Remote Sens.* **2017**, *38*, 558–571. [CrossRef]

11. Dembélé, M.; Zwart, S.J. Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. *Int. J. Remote Sens.* **2016**, *37*, 3995–4014. [CrossRef]

12. Mondal, A.; Lakshmi, V.; Hashemi, H. Intercomparison of trend analysis of Multisatellite Monthly Precipitation Products and Gauge Measurements for River Basins of India. *J. Hydrol.* **2018**, *565*, 779–790. [CrossRef]

13. 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]

14. Wang, Z.; Zhong, R.; Lai, C.; Chen, J. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.* **2017**, *196*, 151–163. [CrossRef]

15. Beria, H.; Nanda, T.; Bisht, D.S.; Chatterjee, C. Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6117–6134. [CrossRef]

16. Yuan, F.; Zhang, L.; Wah, K.; Win, W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Assessment of GPM and TRMM Multi-Satellite Precipitation Products in Streamflow Simulations in a Data-Sparse Mountainous Watershed in Myanmar. *Remote Sens.* **2017**, *9*, 302. [CrossRef]

17. Chen, C.; Chen, Q.; Duan, Z.; Zhang, J.; Mo, K.; Li, Z.; Tang, G. Multiscale comparative evaluation of the GPM IMERG v5 and TRMM 3B42 v7 precipitation products from 2015 to 2017 over a climate transition area of China. *Remote Sens.* **2018**, *10*, 944. [CrossRef]

18. Wang, S.; Liu, J.; Wang, J.; Qiao, X.; Zhang, J. Evaluation of GPM IMERG V05B and TRMM 3B42V7 Precipitation Products over High Mountainous Tributaries in Lhasa with Dense Rain Gauges. *Remote Sens.* **2019**, *11*, 2080. [CrossRef]

19. Wu, Y.; Zhang, Z.; Huang, Y.; Jin, Q.; Chen, X.; Chang, J. Evaluation of the GPM IMERG v5 and TRMM 3B42 v7 Precipitation Products in the Yangtze River Basin, China. *Water* **2019**, *11*, 1459. [CrossRef]

20. Tang, G.; Long, D.; Hong, Y.; Gao, J.; Wan, W. Documentation of multifactorial relationships between precipitation and topography of the Tibetan Plateau using spaceborne precipitation radars. *Remote Sens. Environ.* **2018**, *208*, 82–96. [CrossRef]

21. Jiang, L.; Bauer-Gottwein, P. How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. *J. Hydrol.* **2019**, *572*, 486–500. [CrossRef]

22. Zhang, Z.; Tian, J.; Huang, Y.; Chen, X.; Chen, S. Hydrologic Evaluation of TRMM and GPM IMERG Satellite-Based Precipitation in a Humid Basin of China. *Remote Sens.* **2019**, 431. [CrossRef]

23. Mayor, Y.G.; Tereshchenko, I.; Fonseca-Hernández, M.; Pantoja, D.A.; Montes, J.M. Evaluation of error in IMERG precipitation estimates under different topographic conditions and temporal scales over Mexico. *Remote Sens.* **2017**, *9*, 503. [CrossRef]

24. Sungmin, O.; Foelsche, U.; Kirchengast, G.; Fuchsberger, J.; Tan, J.; Petersen, W.A. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6559–6572. [CrossRef]

25. Huang, W.R.; Chang, Y.H.; Liu, P.Y. Assessment of IMERG precipitation over Taiwan at multiple timescales. *Atmos. Res.* **2018**, *214*, 239–249. [CrossRef]

26. Boluwade, A.; Stadnyk, T.; Fortin, V.; Roy, G. Assimilation of precipitation Estimates from the Integrated Multisatellite Retrievals for GPM (IMERG, early Run) in the Canadian Precipitation Analysis (CaPA). *J. Hydrol. Reg. Stud.* **2017**, *14*, 10–22. [CrossRef]

27. Nashwan, M.S.; Shahid, S. Assessment of Satellite-Based Precipitation Measurement Products over the Hot Desert Climate of Egypt. *Remote Sens.* **2019**, 555. [CrossRef]

28. Anjum, M.N.; Ahmad, I.; Ding, Y.; Shangguan, D.; Zaman, M.; Ijaz, M.W.; Sarwar, K.; Han, H.; Yang, M. Assessment of IMERG-V06 Precipitation Product over Di ff erent Hydro-Climatic Regimes in the Tianshan. *Remote Sens.* **2019**, *11*, 2314. [CrossRef]

29. Lu, M.A. Evaluation of the integrated multi-satellite retrievals for global precipitation measurement over the Tibetan Plateau. *J. Mt. Sci.* **2019**, *16*, 1500–1514.

30. Turini, N.; Thies, B.; Bendix, J. Estimating High Spatio-Temporal Resolution Rainfall from MSG1 and GPM IMERG Based on Machine Learning: Case Study of Iran. *Remote Sens.* **2019**, *11*, 2307. [CrossRef]

31. 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]

32. Ghorbani, M.; Deposits, M.; Resources, N. *The Economic Geology of Iran*; Springer: Dordrecht, the Netherlands, 2013; ISBN 9789400756243.

33. Balling, R.C.; Keikhosravi Kiany, M.S.; Sen Roy, S.; Khoshhal, J. Trends in Extreme Precipitation Indices in Iran: 1951–2007. *Adv. Meteorol.* **2016**, *2016*. [CrossRef]

34. Alizadeh-Choobari, O.; Najafi, M.S. Extreme weather events in Iran under a changing climate. *Clim. Dyn.* **2018**, *50*, 249–260. [CrossRef]

35. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. *IMERG Tech. Doc.* **2017**, 1–60. [CrossRef]

36. Hashemi, H.; Nordin, M.; Lakshmi, V.; Huffman, G.J.; Knight, R. Bias correction of long-term satellite monthly precipitation product (TRMM 3B43) over the conterminous United States. *J. Hydrometeorol.* **2017**, *18*, 2491–2509. [CrossRef]

37. Hosseini, S.H. Disastrous floods after prolonged droughts have challenged Iran. *FUF Bladet* **2019**, *2019*, 30–33.

38. Hashemi, H.; Berndtsson, R.; Persson, M. Floodwater harvesting for artificial recharge and spate irrigation in arid area. *J. Water Manag. Res.* **2013**, *69*, 93–100.

39. Tan, M.L.; Santo, H. Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia. *Atmos. Res.* **2018**, *202*, 63–76. [CrossRef]

Article

Assessment of Precipitation Estimation from the NWP Models and Satellite Products for the Spring 2019 Severe Floods in Iran

Saleh Aminyavari [1], Bahram Saghafian [2,*] and Ehsan Sharifi [3,4]

[1] Department of Civil Engineering, Chalous Branch, Islamic Azad University, Chalous 46615/397, Iran;
    syavari@iauc.ac.ir
[2] Department of Civil Engineering, Science and Research Branch, Islamic Azad University,
    Tehran 1477893855, Iran
[3] Institute of Meteorology and Climatology, University of Natural Resources and Life Sciences (BOKU),
    Vienna 1180, Austria; ehsan.sharifi@boku.ac.at
[4] Department of Meteorology and Geophysics, University of Vienna, Vienna 1090, Austria
[*] Correspondence: b.saghafian@gmail.com or b.saghafian@srbiau.ac.ir

Abstract: Precipitation monitoring and early warning systems are required to reduce negative flood impacts. In this study, the performance of ensemble precipitation forecasts of three numerical weather prediction (NWP) models within the THORPEX interactive grand global ensemble (TIGGE) as well as the integrated multi-satellite retrievals for global precipitation measurement (GPM), namely IMERG, for precipitation estimates were evaluated in recent severe floods in Iran over the March–April 2019 period. The evaluations were conducted in three aspects: spatial distribution of precipitation, mean areal precipitation in three major basins hard hit by the floods, and the dichotomous evaluation in four precipitation thresholds (25, 50, 75, and 100 mm per day). The results showed that the United Kingdom Met Office (UKMO) model, in terms of spatial coverage and satellite estimates as well as the precipitation amount, were closer to the observations. Moreover, with regard to mean precipitation at the basin scale, UKMO and European Center for Medium-Range Weather Forecasts (ECMWF) models in the Gorganrud Basin, ECMWF in the Karkheh Basin and UKMO in the Karun Basin performed better than others in flood forecasting. The National Centers for Environmental Forecast (NCEP) model performed well at low precipitation thresholds, while at high thresholds, its performance decreased significantly. On the contrary, the accuracy of IMERG improved when the precipitation threshold increased. The UKMO had better forecasts than the other models at the 100 mm/day precipitation threshold, whereas the ECMWF had acceptable forecasts in all thresholds and was able to forecast precipitation events with a lower false alarm ratio and better detection when compared to other models.

Keywords: TIGGE; precipitation; numerical weather prediction; satellite; IMERG; flood; spring 2019; Iran

1. Introduction

In recent decades, observed climate trends have shown an increase in temperature worldwide so that extreme precipitation has increased in some specific areas (e.g., eastern half of North America, Eastern Europe, Asia, and South America) [1–5]. Due to rising temperatures, particularly over the Arctic, the sea-ice retreat allows for increased transport of heat and momentum from the ocean up to the tropo- and stratosphere. In the upper atmosphere, these waves deposit the momentum transported, disturbing the stratospheric polar vortex, which can lead to a breakdown of this circulation with the potential to also significantly impact the troposphere in mid- to late-winter and early spring [6,7].

Iran's climate is generally semi-arid and is subject to frequent flooding, causing major damage to people and society. In spring 2019, major floods occurred almost concurrently in different parts of the country. The first flood event occurred in late winter to early spring 2019 in the northeastern provinces due to heavy precipitation over the March 17–22 period. In at least one station, over 280 mm of precipitation was recorded over the six-day period. The subsequent second and third flood events occurred in the March 24–26 and March 31–April 2 periods, respectively, where most of the precipitation fell in the southwest, causing widespread damage to the people and infrastructure while filling/causing an overflow of most reservoirs. The total economic cost of these floods is estimated to be $3.5 billion U.S. dollars. Studies on the causes of the March–April 2019 severe floods are still ongoing, although exceptional precipitation and climate change attribution are on the minds of most experts.

Given the heavy negative impacts imposed by the 2019 flood events, the monitoring and forecasting of precipitation remain major challenges for hydrologists and reservoir managers. The availability of global ensemble forecast models in the THORPEX Interactive Grand Global Ensemble (TIGGE) database [8] as well as high-resolution satellite estimates creates new opportunities for flood monitoring/forecasting. Extensive research has been conducted on the application of satellite precipitation estimates for motoring purposes and numerical weather prediction (NWP) models for forecast objectives. In terms of the latter, using TIGGE forecasts for a flood alarm system in China [9], flood early warning with European Centre for Medium-Range Weather Forecasts (ECMWF) model forecasts under Global Flood Awareness System (GLoFAS) in global scale, European Flood Awareness System (EFAS) for Europe projects [10], and African Flood Forecasting System (AFFS) [11] are some examples of NWP applications in flood forecasting.

Numerous studies have been conducted to evaluate the estimated precipitation from NWP models and satellite-based precipitation estimates (SPEs). For example, the results of TIGGE precipitation forecasting in flood-prone areas of China showed that the ensemble forecast model is more proficient than the single forecast models [12]. He et al. [13] showed that TIGGE ensemble forecasts are suitable for forecasting flood events. The rapid alert system was developed from four operational NWP models: UKMO (United Kingdom Met Office), NCEP (National Centers for Environmental Forecast), ECMWF, and JMA (Japan Meteorological Agency). The probability of severe weather events was forecasted based on the climatological probability density function in each model. Numerous case studies have shown that these products successfully forecasted severe events such as the Russian heat wave in 2010, the Pakistan flood in 2010, and Hurricane Sandy in 2012 [14]. In West Africa, evaluating the forecasts of seven TIGGE meteorological databases against the Tropical Precipitation Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) product assumed as the observed precipitation over the 2008 to 2012 period showed that ECMWF and UKMO performed better than the other models [15]. In Iran, the forecasts of ECMWF, UKMO, and NCEP centers for thirteen synoptic stations in eight different precipitation regions over a 1–3 day lead time over the 2008 to 2016 period showed that ECMWF in most regions, UKMO in mountainous areas, and NCEP along the Persian Gulf coast performed the best, while, as expected, the model skill decreased with increasing lead time [16].

The accuracy of SPEs is influenced by their spatiotemporal resolution, which in turn impacts the prediction of natural hazards. Therefore, the assessment of new precipitation products is often recommended before the product can be employed in research and decision-making. Several studies have been conducted to evaluate the SPEs over Iran. Moazzami et al. [17] and Javanmard et al. [18] examined different SPEs products at a daily time-scale over diverse climate conditions in Iran. Overall, the results showed that 3B42V7 outperformed other SPEs. Sharifi et al. [19] evaluated the first version of the integrated multi-satellite retrievals for global precipitation measurement (IMERG) version-03 in comparison with the TMPA and ERA-Interim products across different parts of Iran and found that IMERG generally outperformed the other products. They later improved the spatial resolution and accuracy of the SPEs through downscaling and bias-correction techniques [20,21]. In another study, Beck et al. [22] evaluated 26 precipitation datasets and compared them with gauge-radar data over the United States. Among the gauge-corrected products, the best overall performance was obtained

by Multi-Source Weighted-Ensemble Precipitation (MSWEP)-V2.2, followed by IMERGDF-V05 and MERRA-2. However, IMERG real-time V05 performed substantially better than TMPA-3B42-real-time V7 and ERA5-HRES, particularly, in regions dominated by convective storms. In another study, Sharifi et al. [23] examined the accuracy of six SPEs and gridded precipitation models against a dense network of 872 stations over Austria, in terms of extreme events and different altitude categories. They also found that the latest version of IMERG-V06A performed better than the other products (except MSWEP-2.2), which was consistent with the study results by Beck et al. [22]. With respect to extreme precipitation events, Fang et al. indicated that although IMERG well captured the spatial pattern of extreme precipitation over China, the topography and climate condition had a significant influence on its performance [24]. In another study, Sunikumar et al. demonstrated the ability of IMERG to follow the intraseasonal variability with minor differences observed in the maximum values of precipitation during the rainy season over Japan, Philippines, and Nepal [25]. Mazzoglio et al. improved an extreme precipitation detection system using IMERG data and stated that this product guarantees good results when the precipitation aggregation interval is equal to or greater than 12 h [26]. However, no comparison has been made so far between the ensemble forecasts with IMERG product for extreme flood events over Iran.

The purpose of this study was to evaluate the performance of ensemble precipitation forecasting models and IMERG products for three severe 2019 flood events in Iran to determine whether these products have potential in (major) flood warning applications over Iran. According to ground precipitation data, the highest precipitation occurred in Gorganrud, Karkheh, and Karun Basins. These three basins constitute the study area.

2. Data and Methods

Daily in situ accumulated precipitation data from 100 meteorological synoptic stations and 221 Ministry of Energy rain gauges between 15 March 2019 to 2 April 2019 were collected. The collected data have been quality controlled by the respective organization. Ensemble precipitation forecasts of three major global meteorological centers, namely ECMWF, NCEP and UKMO, were then extracted from the TIGGE database in a 24-h lead time (https://confluence.ecmwf.int/display/TIGGE). TIGGE is part of the THORPEX project, which includes ensemble forecasts of 11 NWP world centers. Ensemble forecasts include multiple individual forecasts generated by different physical parameterizations or different initial conditions. Based on the results of NWP evaluation in several studies, the three selected models performed better than the others in different regions of Iran/the world [9,27]. As a result, the NCEP, UKMO, and ECMWF models were evaluated as numerical forecasts in this study. The NWP data are in the GRIB2 format with a resolution of 50 km. Furthermore, daily IMERG-V06B-RT satellite estimates with an approximate 10-km spatial resolution for the whole of Iran were downloaded in NetCDF format for the study time period. Figure 1 shows the spatial distribution of the in situ observations overlaid on the elevation map. Table 1 shows the characteristics of the three NWP models and IMERG product.
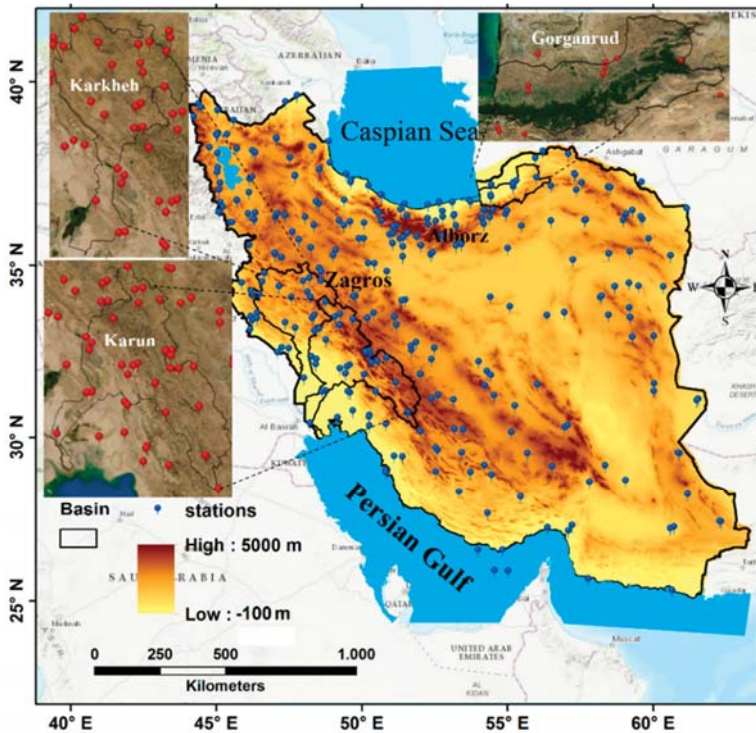
Figure 1. The spatial distribution of in situ observations on the elevation map of Iran.

Table 1. Characteristics of the studied NWP models and IMERG satellite.

Products	Base Time (UTC)	No. of Ensemble Members	Spatial Resolution	Temporal Resolution	Forecast Length (day)
ECMWF	00/12	50 + 1	~0.28° × (~0.56°)	6 h	0–15
NCEP	00/06/12/18	20 + 1	1.0° × 1.0°	6 h	0–16
UKMO	00/12	17 + 1	0.83° × 0.56°	6 h	0–15
IMERG	—	—	0.1° × 0.1°	30 min	—

In this study, the evaluations were conducted in three steps. In the first step, the performance of the three NWP models and SPEs in terms of capturing the spatial distribution of precipitation for the three flood events (17–22 March, 24–26 March, and 31 March to 2 April 2019, respectively) were compared. The numerical forecast and SPE data have a 50 × 50 km and ~10 × 10 km spatial resolution, respectively. Therefore, spatial aggregation from 10 km to 50 km was performed using the cubic convolution resampling method, which is based on the weighted average of 16 nearest neighboring pixels [28]. An in situ observation map was also constructed using inverse distance weighting (IDW) interpolation with a 50 × 50 km resolution. Moreover, for a more robust comparison of the spatial distribution of precipitation, the isohyet contours over the three respective basins for all three flood events were obtained.

For detailed insight into the second step, the mean, maximum, and minimum precipitation values for each individual event were also determined.

In the third step, the NWP forecasts and satellite data were interpolated with the aim of a direct comparison of precipitation with the in situ measurements. The IDW method used for interpolation involved four grid-points around each station. Then the interpolated precipitation at each station was

evaluated against the observation. The dichotomous (yes/no) evaluation of daily precipitation was further examined. For this purpose, 25, 50, 75, and 100 mm/day thresholds were set and the number of correct events determined by the satellite and NWP models in each threshold were compared. Precipitation events were counted if at least one of the stations operating in each basin recorded precipitation. Otherwise, if none of the stations recorded precipitation, a "no-precipitation" event was assigned to the whole basin. Accordingly, the probability of detection (POD) and false alarm ratio (FAR) as well as the equitable threat score (ETS) criterion were used to examine the capability of the products to detect the precipitation events. POD and ETS values vary between 0 to 1, with 1 as a perfect score, while the FAR perfect score is 0. In addition, the average results of the dichotomous evaluation of the stations in each basin were calculated in four precipitation thresholds. Interested readers are referred to Wilks (2011) for further detail on the dichotomous (yes/no) evaluation [29]. Table 2 indicates the metrics used to measure the effectiveness of precipitation estimations.

Table 2. Summary of the assessment criteria used in this study [27].

Verification Measure	Formula	Description	Perfect/no Skill
Pearson's correlation coefficient	$CC = \dfrac{\sum(F-\bar{F})(O-\bar{O})}{\sqrt{\sum(F-\bar{F})^2}\sqrt{\sum(O-\bar{O})^2}}$	Linear dependency between forecast and observation	1/0
Probability of detection (Hit Rate)	POD = A/(A + C)	What fraction of the observed "yes" events were correctly forecasted?	1/0
False alarm ratio	FAR = B/(B + C)	What fraction of the predicted "yes" events actually did not occur	0/1
Equitable threat score	ETS = (A − A$_{random}$)/(A + B + C − A$_{random}$) A$_{random}$ = (A + C)(A + B)/N	How well did the forecasted "yes" events correspond to the observed "yes" events?	1/0

Notes: *F* and *O* denote the forecast and corresponding observation, respectively. Similarly, $\bar{F}$ and $\bar{O}$ denote the forecast average and observation average, respectively. A, B, C, and D were obtained from the contingency table.

It should be noted that the evaluations of the first and second step were based on the total precipitation in each flood event. To clarify this, the first precipitation/flood event consisted of the precipitation accumulation of six days from 17 March to 22 March 2019; the second event in three days from 24 March to 26 March 2019; and the third event in three days from 31 March to 2 April 2019. A tertiary evaluation for the daily accumulation precipitations from 15 March to 2 April 2019 was conducted.

3. Results

By presenting the results, this section provides insights into where the model/satellite systematically differ and discusses the possible causes.

Spatial distribution of precipitation accumulation during the flood days (15 March to 2 April 2019) corresponding to the in situ observations, satellite, and three NWP models are shown in Figure 2. It is clear that high precipitation areas extended from the northeast along the Alborz Mountains to the southeastern Caspian Sea; and from the west to southwest along the Zagros Mountains. Overall, the precipitation spatial patterns in all products were correctly captured, although there were differences in the precipitation magnitude among the products. In more specific terms, IMERG precipitation product (IMERG-RT V05) reasonably captured the precipitation distribution for most parts of the country when compared with the in situ observations. Moreover, the in situ observations highlighted four precipitation hotspots along the Zagros Mountain in the west of Iran with the largest precipitation amount. IMERG outperformed the other products to capture these hotspots. The remarkable precipitation gradients were well-captured by IMERG, possibly due to its native higher spatial IMERG resolution (~10 km × ~10 km) when compared with those of the NWP models (50 km × 50 km). Among the NWP models, the NCEP model forecasted much lower precipitation, while

UKMO generally overestimated the in situ observations. In general, most of the precipitation products could capture the basic precipitation characteristics in terms of the observed spatial pattern, however, it is unknown as to whether the overestimation/underestimation was really derived by the product structure or if it was due to lack of dense in situ observations. It should be noted that the mean of the ensemble NWP forecasts was used in Figure 2.
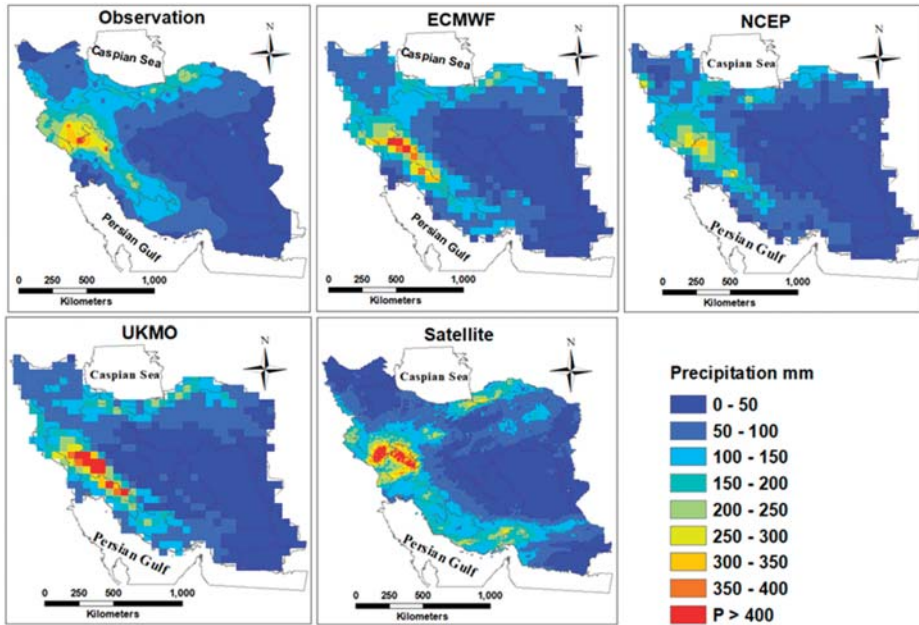
Figure 2. Spatial distribution of accumulated precipitation (mm) from the observational, ECMWF, NCEP, UKMO, and IMERG satellite for 19 precipitation days over Iran.

3.1. Spatial Distribution of Precipitation

In this section, the differences between the in situ observations and NWP/satellite are shown with a 50 km spatial resolution in Figure 3. In the first flood event in northeastern Iran, the UKMO model outperformed other models and indicated lower precipitation differences with the observations. Moreover, the NCEP underestimated the extreme precipitation over the southeast of the Caspian Sea, while IMERG displayed differences over the southeast and no difference over the southwest part of the Caspian Sea. In the second flood event in western and southwestern Iran, the ECMWF outperformed other products. The NCEP performed better in this flood event than it did in the first flood event. However, all models/satellite, especially UKMO, generally overestimated the precipitation amount. In the third flood event in western Iran, IMERG reasonably captured the spatial precipitation for most parts of the region, except in a few pixels in the west. In addition, IMERG indicated a smoother trend in precipitation differences from the northwest to the eastern part of Iran, while it is evident that other products had larger errors in terms of the difference with the in situ observations.
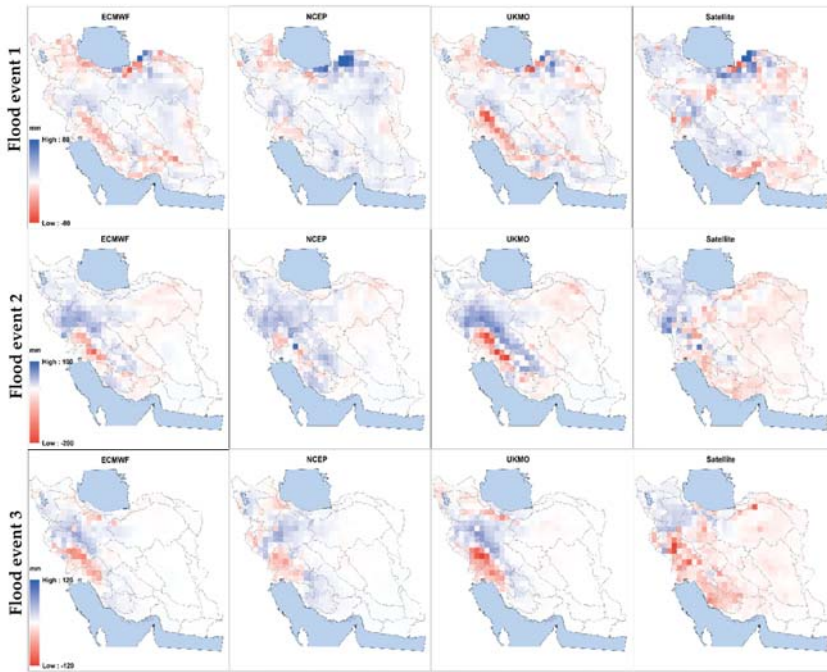
Figure 3. Differences between the accumulated precipitation in situ observations, NWP, and IMERG satellite data.

Table 3 presents the quantitative evaluation of precipitation products. Based on this table, during the first flood event, ECMWF and IMERG were closer to the in situ observation data with respect to the Correlation Coefficient (CC) and maximum precipitation amount. Although the NCEP showed the least correlation in the first event, this model indicated relatively close forecasts to the observations for the second and third events. In general, the ECMWF showed slightly better skill in terms of the correlation, min, and max values of precipitation in comparison with the other products (Table 3).

Table 3. Statistical characteristics of the observations, NWP, and IMERG satellite precipitation data.

		MEAN (mm)	MAX (mm)	STD (mm)	Correlation
	Observation	18.38	237	21.1	—
	ECMWF	18.24	168	20.84	0.8
Flood event 1	UKMO	18.02	177	22.1	0.77
	NCEP	13.65	101	12.2	0.73
	IMERG	15.34	191	20.68	0.73
	Observation	24.64	208	27.3	—
	ECMWF	25.12	239	29.62	0.82
Flood event 2	UKMO	26.13	298	38.16	0.72
	NCEP	21.85	161	24.37	0.85
	IMERG	31.6	240	30.73	0.79
	Observation	15.15	197	26.14	—
	ECMWF	14.73	173	25.88	0.87
Flood event 3	UKMO	15.53	194	27.13	0.78
	NCEP	13.39	113	23.42	0.87
	IMERG	20.67	298	31.77	0.85

For a more in-depth evaluation of the spatial pattern, the isohyet contours indicate that for the first flood event, which mostly impacted the Gorganrud Basin in northeast Iran, the UKMO and ECMWF models performed similarly, whereas the UKMO model better forecasted the location of heavy precipitation, followed by the satellite (IMERG). The NCEP model, although recognizing the location of heavy precipitation, underestimated the amount of heavy precipitation (Figure 4).
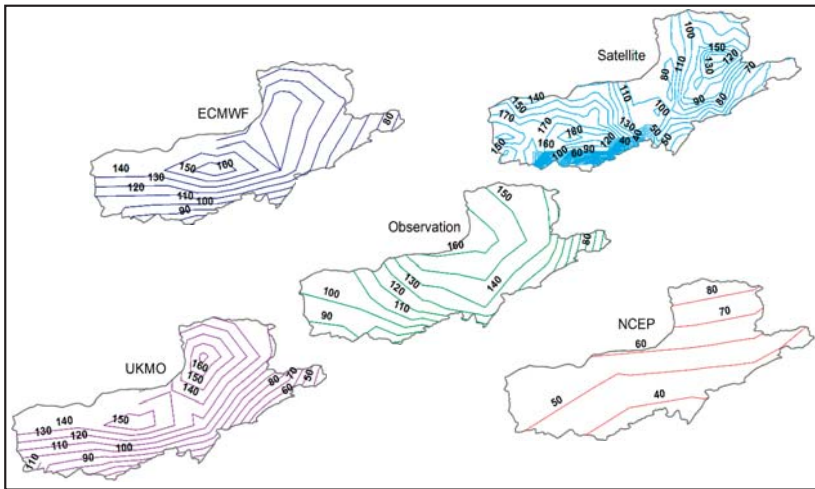
Figure 4. Isohyet contours in the Gorganrud Basin during the first flood event.

According to Figure 5 for the second flood event, the UKMO in the Karkheh Basin yielded better results in comparison with other products in the detection of heavy precipitation. In the Karun Basin, although the UKMO and IMERG products overestimated precipitation, they generally demonstrated better skills than the other two NWP models. The NCEP model in the southeast of the Karun Basin better forecasted heavy precipitation than the other products. With respect to the third flood event in the Karkheh Basin, the UKMO and IMERG showed better performance in the detection of the location of heavy precipitation. In the Karun Basin, although all products showed acceptable performance, the UKMO and ECMWF models outperformed the other products (Figure 6).
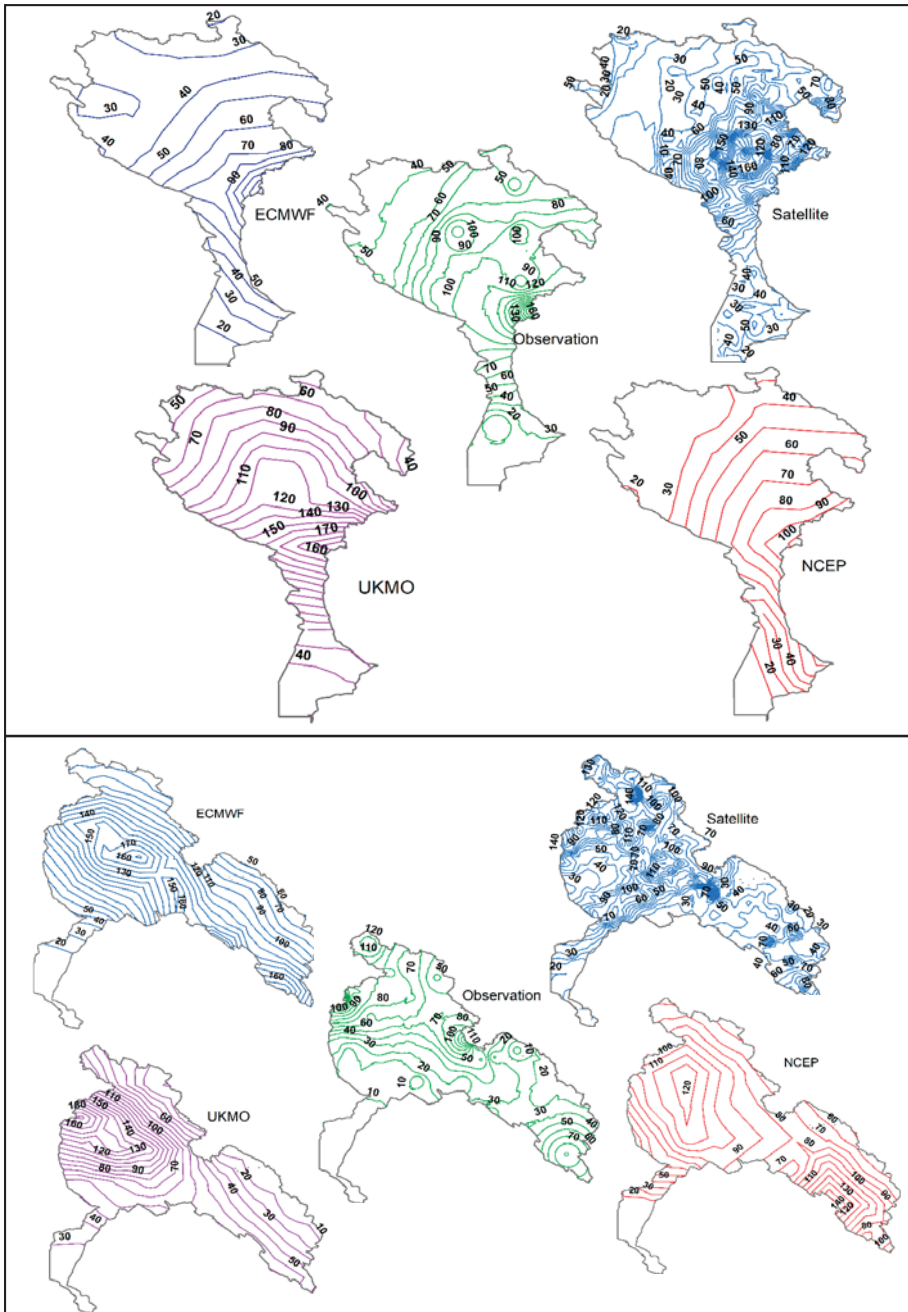
Figure 5. Isohyet contours in the Karkheh (**top**) and Karun (**bottom**) Basins during the second flood event.
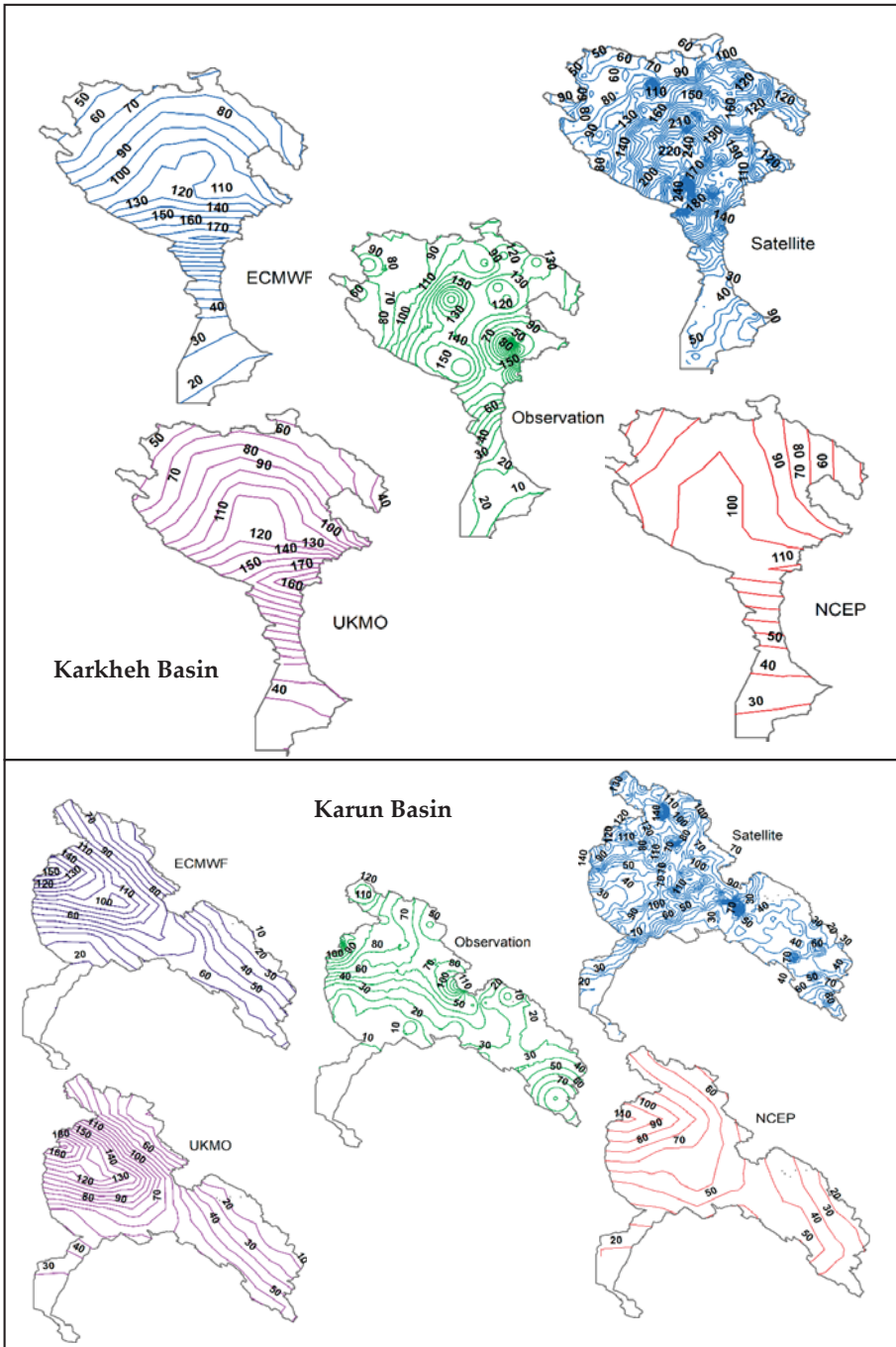
Figure 6. Isohyet contours in Karkheh (**top**) and the Karun (**bottom**) Basins during the third flood event.

3.2. Mean, Minimum and Maximum Basin Precipitation

In order to compare the performances of the precipitation products in each individual basin within the three flood events that occurred from March to April 2019, the mean, minimum, and maximum precipitations were plotted, as shown in Figure 7. The first flood event occurred in the Gorganrud Basin in northeastern Iran. As Figure 7a shows, in terms of mean, maximum, and minimum precipitation, the ECMWF and UKMO models performed approximately the same and well forecasted the min, max, and mean of precipitation. The NCEP model also forecasted less precipitation than the observations with respect to mean precipitation. However, the maximum in situ observations in this basin was higher than all the NWP and IMERG satellite products. Although all evaluated precipitation products were rather close to the in situ measurements, it is evident that ECMWF and UKMO, followed by IMERG, compared well with the corresponding in situ measurements in terms of mean precipitation through the first event.

Figure 7b shows the mean, min, and max value of precipitation obtained from all products and in situ observations over the southwest of Iran within the Karkheh and Karun Basins. The mean precipitation values indicated that generally all of the models and satellite well estimated the mean areal precipitation with a slight underestimation during this event over the Karkheh Basin. However, in terms of maximum precipitation, the UKMO overestimated, while the satellite and other models showed underestimation over the Karkheh Basin during the second flood event. This might be due to high moisture in the atmosphere and the occurrence of small scale precipitation systems that were dominant during this event [23]. Additionally, from the median and the 25th and 75th percentiles of the box-plots, one can infer that the precipitation estimated by IMERG, followed by ECMWF, was more accurate than the other products in both basins, while the UKMO whiskers extended to the highest extreme data points in comparison to the other products.

According to Figure 7c, the mean areal precipitation values of all products were rather close to the in situ observations over the Karkheh Basin, while the box-plots confirmed that the IMERG and NCEP values were in a larger and smaller precipitation ranges in comparison to the in situ observations, respectively. However, in the Karun Basin, the UKMO greatly overestimated the precipitation, while IMERG, followed by NCEP, showed improvement in heavy daily precipitation in comparison with the in situ observations.
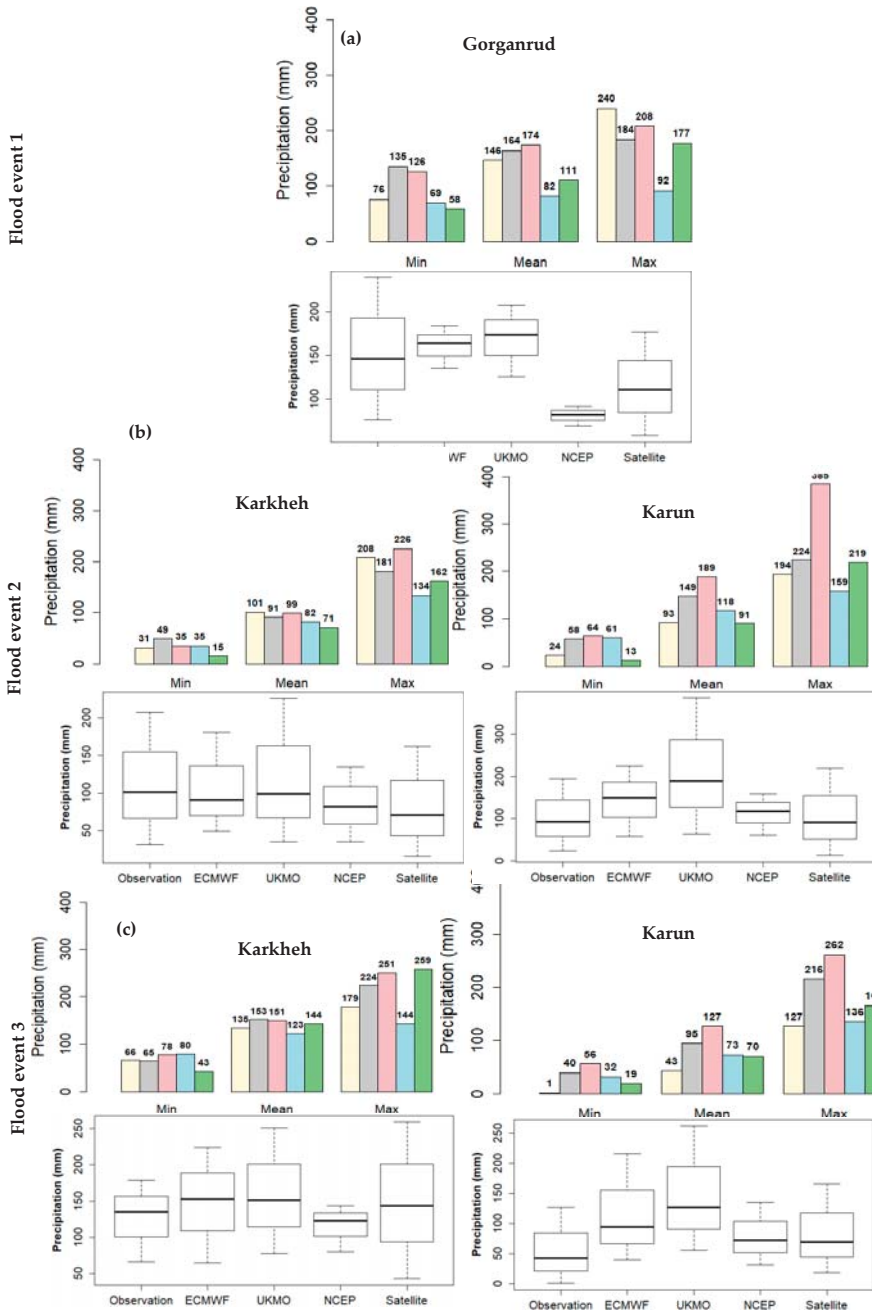
Figure 7. Mean, maximum, minimum precipitation amounts and box-plots in the (**a**) Gorganrud, Karkheh, and Karun Basins for the a) first, (**b**) second, and (**c**) third flood events. The center-line of each boxplot depicts the median value (50th percentile) and the box encompasses the 25th and 75th percentiles of the sample data, while the whiskers represent the extreme values, respectively.

3.3. Dichotomous Evaluation

With respect to dichotomous evaluation, Figure 8 compares the number of 24-h extreme precipitation events recorded at all stations in different precipitation thresholds. As the main emphasis of this study was on the evaluation of extreme events, the maximum of ensemble forecasts was carefully determined. Based on Figure 8, at a threshold of 25 mm, NWP models estimated precipitation better than IMERG, while among the NWP models, ECMWF was closer to in situ observations. However, all products underestimated the number of events above 25 mm. At the 50 mm threshold, the UKMO predictions of the number of events were closer to the observations. However, in terms of mean ensemble forecast, IMERG estimated the number of events closer to the observations than the NWP models. It is worth noting that the performance of the NCEP model decreased significantly with an increase in the threshold. At the 75 mm threshold, UKMO was better than the other products, whereas NCEP had difficulty in forecasting the precipitation amount at this threshold. From 15 March 2019 to 2 April 2019, precipitation amounts of over 100 mm were reported for a number of days. At this threshold, in situ observations recorded 22 events, while NCEP detected none. However, the UKMO and ECMWF maximum ensembles as well as IMERG detected 11, 10, and 4 events, respectively.
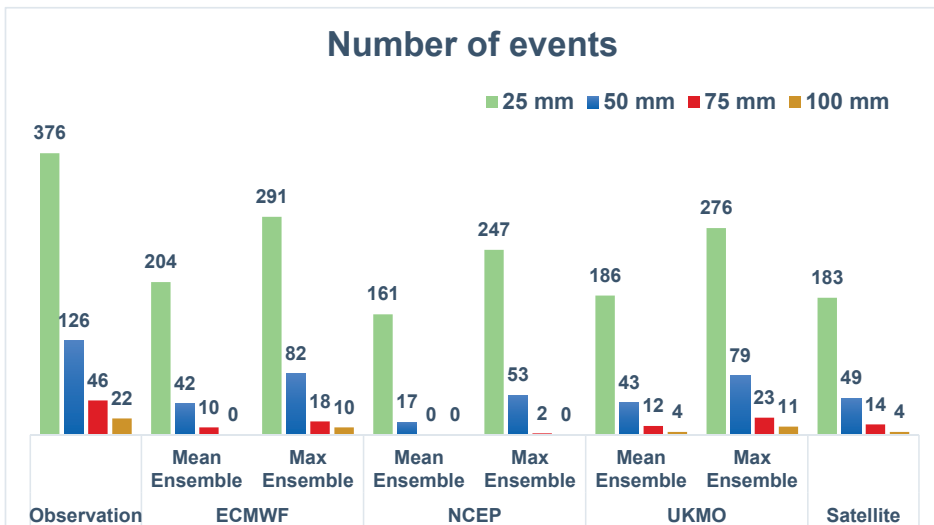
Figure 8. Comparison of the number of extreme events estimated by the NWP models and IMERG satellite with observations in different precipitation thresholds.

The performance of all precipitation products was then evaluated using dichotomous evaluation criteria (Figure 9). Based on the POD metric, which measures the percentage of the accurate detection of events, the ECMWF in the 25 mm threshold and the UKMO in 50, 75, and 100 mm thresholds showed the highest PODs over Iran. The decline in the performance of NCEP with an increase in the threshold is clearly evident in this criterion. However, based on the FAR metric, which expresses the percentage of false alarms, ECMWF obtained the lowest FAR in maximum ensemble forecasts within all thresholds, although it did not perform well in high thresholds in the mean ensemble forecast mode. Based on the ETS, which measures the forecast quality of detecting occurrence and non-occurrence of an event, the mean ensemble forecast of the ECMWF model achieved the best score in the 25 mm threshold. In the other thresholds, the maximum ensemble forecast of the ECMWF model achieved the highest ETS scores.
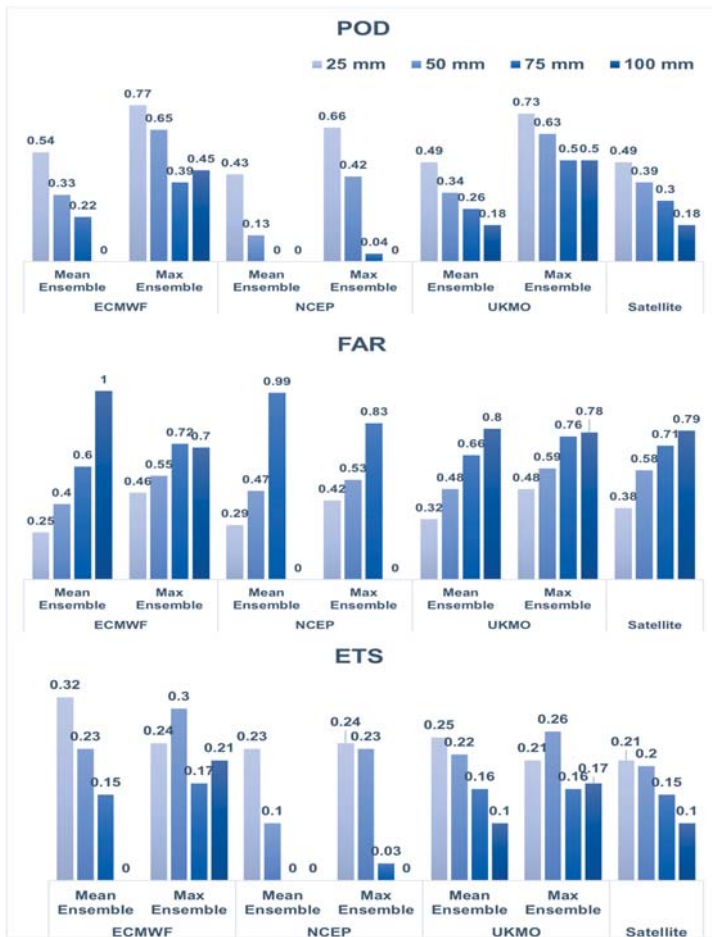
Figure 9. Verification statistics (POD, FAR, and ETS) between ECMWF, NCEP, UKMO, IMERG, and the in situ observations for different thresholds over Iran for 15 March 2019 to 2 April 2019.

Finally, the evaluation with respect to the number of events estimated by the precipitation products at different thresholds was conducted. According to Figure 10, in Gorganrud basin in 25 mm threshold, both UKMO and ECMWF models performed well. Satellite (IMERG) and NCEP obtained rather similar performances with the detection of seven events among the 13 events in these thresholds, while the ECMWF and UKMO performances were more robust and closer to the in situ observations with the detection of 12 events. In the 50 mm threshold, both ECMWF and UKMO still performed better than the other products with the detection of seven events among the 11 events detected by the in situ observations, while the skill of NCEP declined and IMERG showed an improvement with the detection of two and six events, respectively. At a threshold of 75 mm, both ECMWF and UKMO models performed well and could forecast almost 80% of events (five out of seven events) in the 75 mm threshold. The NCEP model did not record any event and IMERG estimated only one event out of seven events. In the 100 mm threshold, both ECMWF and UKMO forecasted five events out of six events above 100 mm. The satellite and NCEP model were weak in this segment and did not estimate any events in northeast Iran in the Gorganroud Basin.
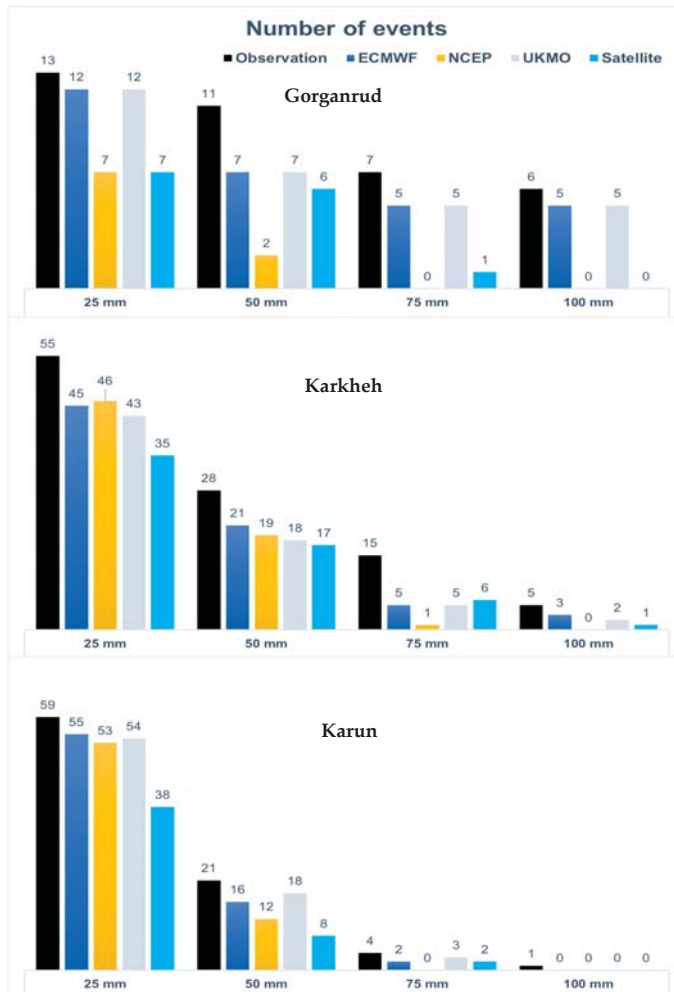
Figure 10. Comparison of the number of extreme events estimated by ECMWF, NCEP, UKMO, and IMERG with in situ observations.

In the Karkheh Basin, the NCEP model performed well in the 25 mm threshold and was able to estimate the number of events better than the other products. However, the performance of NCEP decreased dramatically with an increase in the threshold (e.g., in the 75 mm (100 mm) threshold only one (zero) event out of 15 (5) events were forecasted). Meanwhile, IMERG performed better at higher thresholds. Generally, the ECMWF revealed better results than the other products in all precipitation categories.

In Karun Basin, NCEP again performed better at lower thresholds than at higher thresholds, whereas UKMO was slightly better than ECMWF. However, extreme precipitation greater than 100 mm was only recorded once in this basin.

4. Discussion

In this study, the evaluation of three precipitation forecast models, namely ECMWF, NCEP, and UKMO as well as the IMERG-RT V05B satellite product provided new insight into how errors vary with extreme precipitation events within different climate zones of Iran. Overall, the examined products in this study sometimes agreed well with in situ observations while in some other instances, showed significant differences. As far as possible causes for model performances, precipitation is a function of available atmospheric moisture while derived from moisture convergence. Thus, the models need to be correctly initialized and parameterized through several factors such as (i) gross condensation rate, (ii) latent heat energy exchange within the atmosphere, and (iii) the microphysical behavior of clouds [30]. However, individual clouds typically occur at subgrid scales and must be parameterized based on resolved variables such as average humidity and temperature [31]. The parameterization of cloud, and thus precipitation, continues to be one of the greatest sources of uncertainty in NWP models [32].

Another is the significant differences in the detection of extreme precipitation amounts among the products when compared with the in situ observations. Although, in general, the models captured the spatial distribution of heavy precipitation events, the hot spots were not located in the correct area. Moreover, orography and local effects can affect the accuracy of the products. These issues should be addressed by improving the models' algorithms [33].

Another factor in the interpretation of the differences between the model/satellite products and in situ observations might be related to the precipitation thresholds. As such, overestimation or underestimation in each precipitation threshold means that a given precipitation product was not able to estimate/detect precipitation within that particular threshold, while they could have estimated/detected the precipitation within a lower or higher threshold.

5. Conclusions

In this study, the performance of ensemble precipitation forecasts of three NWP models within the TIGGE database, namely ECMWF, NCEP, and UKMO, and a satellite-based precipitation product, IMERG, for three severe flood events in Iran in March–April 2019 period were evaluated. In the first step, the performance of the precipitation products in capturing the spatial distribution of precipitation was evaluated. The results showed that all of the products could generally capture the main features of the precipitation system, including the spatial distribution, total accumulation, and extreme values (Figure 2). In general, UKMO, followed by IMERG and ECMWF, showed better performance than other products in capturing the spatial distribution of the accumulated precipitation during the 19 days of extreme precipitation events over Iran. However, the in situ observations identified four precipitation hotspots along the Zagros Mountain in western Iran with the largest precipitation amount; IMERG outperformed other products to capture these hotspots.

In the second step, all of the products were examined in comparison with the in situ observations in three major basins that were most affected by the floods. It was evident that ECMWF and UKMO, followed by IMERG, compared well with the corresponding in situ measurements in terms of mean precipitation through the first event in the Gorganrud Basin (Figure 7a). With respect to the second flood event, the box-plots indicated that IMERG, followed by ECMWF, outperformed other products in both the Karkheh and Karun Basins, while the UKMO whiskers extended to the most extreme data points (Figure 7b). In the third flood event, the mean areal precipitation values of all products were rather close to in situ observations over the Karkheh Basin, while box-plots confirmed that the IMERG pixels were in a larger range in comparison with the observations. However, almost all products overestimated the precipitation over the Karkheh and Karun Basins (Figure 7c).

In the third evaluation step, four daily precipitation thresholds of 25 mm, 50 mm, 75 mm, and 100 mm were selected to evaluate the skill of the products in capturing precipitation within the specified thresholds via dichotomous evaluation methods. The results showed that when the threshold was increased, the performance of the NCEP model was greatly reduced, while the IMERG estimates

improved at higher thresholds. At the 50 mm threshold, UKMO predicted a closer number of events in comparison with the observations. At the 75 mm threshold, UKMO revealed better results than the other products, whereas NCEP had difficulty in forecasting the precipitation amount at this threshold. At the 100 mm threshold, the in situ observations recorded 22 events, while NCEP detected none. The maximum ensemble forecasts of UKMO in higher thresholds could estimate a larger number of precipitation events than other models and the satellite. As such, UKMO detected 11 out of 22 events (Figure 8). However, in terms of the contingency table, ECMWF outperformed other products with a higher POD and lower FAR (Figure 9).

Overall, the results of this study show that the IMERG precipitation estimates and NWP ensemble forecasts performed well in the three major flood events in spring 2019 in Iran. Given the widespread damage caused by the floods, the necessity of establishing an efficient flood warning system using the best precipitation products is advised.

The overestimation/underestimation of precipitation by forecast models and satellite-based precipitation products still remains a challenge, particularly for extreme precipitation events. Short-time and extreme precipitation events are much more variable than moderate precipitation events. However, studies on the impact of the uncertainty of precipitation products are needed to obtain a better understanding of how and why precipitation products succeed or fail in the detection of heavy precipitation. Moreover, it is important to note that this study was conducted based on a short period of data (i.e., 19 precipitation days containing three severe flood events) limited to Iran. Thus, further studies using a longer dataset at the global scale in different climate regimes/geophysical features is essential to assess the impacts of the aforementioned limitations.

References

1. Wehner, M.F.; Arnold, J.R.; Knutson, T.; Kunkel, K.E.; LeGrande, A.N. Ch. 8: Droughts, Floods, and Wildfires. In *Climate Science Special Report: Fourth National Climate Assessment*; U.S. Global Change Research Program: Washington, DC, USA, 2017; Volume I.

2. Sein, K.K.; Chidthaisong, A.; Oo, A.K.L. Observed Trends and Changes in Temperature and Precipitation Extreme Indices over Myanmar. *Atmosphere* **2018**, *9*, 477. [CrossRef]

3. Alexander, L.V.; Zhang, X.; Peterson, T.C.; Caesar, J.; Gleason, B.; Klein Tank, A.M.G.; Haylock, M.; Collins, D.; Trewin, B.; Rahimzadeh, F.; et al. Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.* **2006**, *111*. [CrossRef]

4. Easterling, D.R. Maximum and Minimum Temperature Trends for the Globe. *Science* **1997**, *277*, 364–367. [CrossRef]

5. Kiktev, D.; Sexton, D.M.H.; Alexander, L.; Folland, C.K. Comparison of Modeled and Observed Trends in Indices of Daily Climate Extremes. *J. Clim.* **2003**, *16*, 3560–3571. [CrossRef]

6. Romanowsky, E.; Handorf, D.; Jaiser, R.; Wohltmann, I.; Dorn, W.; Ukita, J.; Cohen, J.; Dethloff, K.; Rex, M. The role of stratospheric ozone for Arctic-midlatitude linkages. *Sci. Rep.* **2019**, *9*, 7962. [CrossRef] [PubMed]

7. Francis, J.A.; Vavrus, S.J. Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophys. Res. Lett.* **2012**, *39*. [CrossRef]

8. Swinbank, R.; Kyouda, M.; Buchanan, P.; Froude, L.; Hamill, T.M.; Hewson, T.D.; Keller, J.H.; Matsueda, M.; Methven, J.; Pappenberger, F.; et al. The TIGGE Project and Its Achievements. *Bull. Am. Meteorol. Soc.* **2016**, *97*, 49–67. [CrossRef]

9. He, Y.; Wetterhall, F.; Cloke, H.L.; Pappenberger, F.; Wilson, M.; Freer, J.; McGregor, G. Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions. *Meteorol. Appl.* **2009**, *16*, 91–101. [CrossRef]

10. Thielen, J.; Bartholmes, J.; Ramos, M.-H.; Roo, A.D. The European Flood Alert System—Part 1: Concept and development. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 125–140. [CrossRef]

11. Thiemig, V.; Bisselink, B.; Pappenberger, F.; Thielen, J. A pan-African medium-range ensemble flood forecast system. *Hydrol. Earth Sys. Sci.* **2015**, *19*, 3365–3385. [CrossRef]

12. Krishnamurti, T.N.; Sagadevan, A.D.; Chakraborty, A.; Mishra, A.K.; Simon, A. Improving multimodel weather forecast of monsoon rain over China using FSU superensemble. *Adv. Atmos. Sci.* **2009**, *26*, 813–839. [CrossRef]

13. He, Y.; Wetterhall, F.; Bao, H.; Cloke, H.; Li, Z.; Pappenberger, F.; Hu, Y.; Manful, D.; Huang, Y. Ensemble forecasting using TIGGE for the July-September 2008 floods in the Upper Huai catchment: A case study. *Atmos. Sci. Lett.* **2010**, *11*, 132–138. [CrossRef]

14. Matsueda, M.; Nakazawa, T. Early warning products for severe weather events derived from operational medium-range ensemble forecasts. *Meteorol. Appl.* **2015**, *22*, 213–222. [CrossRef]

15. Louvet, S.; Sultan, B.; Janicot, S.; Kamsu-Tamo, P.H.; Ndiaye, O. Evaluation of TIGGE precipitation forecasts over West Africa at intraseasonal timescale. *Clim. Dyn.* **2016**, *47*, 31–47. [CrossRef]

16. Aminyavari, S.; Saghafian, B.; Delavar, M. Evaluation of TIGGE Ensemble Forecasts of Precipitation in Distinct Climate Regions in Iran. *Adv. Atmos. Sci.* **2018**, *35*, 457–468. [CrossRef]

17. Moazami, S.; Golian, S.; Hong, Y.; Sheng, C.; Kavianpour, M.R. Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran. *Hydrol. Sci. J.* **2015**, *61*, 420–440. [CrossRef]

18. Javanmard, S.; Yatagai, A.; Nodzu, M.I.; BodaghJamali, J.; Kawamoto, H. Comparing high-resolution gridded precipitation data with satellite rainfall estimates of TRMM_3B42 over Iran. *Adv. Geosci.* **2010**, *25*, 119–125. [CrossRef]

19. Sharifi, E.; Steinacker, R.; Saghafian, B. Assessment of GPM-IMERG and Other Precipitation Products against Gauge Data under Different Topographic and Climatic Conditions in Iran: Preliminary Results. *Remote Sens.* **2016**, *8*, 135. [CrossRef]

20. Sharifi, E.; Saghafian, B.; Steinacker, R. Downscaling Satellite Precipitation Estimates With Multiple Linear Regression, Artificial Neural Networks, and Spline Interpolation Techniques. *J. Geophys. Res. Atmos.* **2019**, *124*, 789–805. [CrossRef]

21. Sharifi, E.; Saghafian, B.; Steinacker, R. Copula-based stochastic uncertainty analysis of satellite precipitation products. *J. Hydrol.* **2019**, *570*, 739–754. [CrossRef]

22. Beck, H.E.; Wood, E.F.; Pan, M.; Fisher, C.K.; Miralles, D.G.; van Dijk, A.I.J.M.; McVicar, T.R.; Adler, R.F. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bull. Am. Meteorol. Soc.* **2019**, *100*, 473–500. [CrossRef]

23. Sharifi, E.; Eitzinger, J.; Dorigo, W. Performance of the State-Of-The-Art Gridded Precipitation Products over Mountainous Terrain: A Regional Study over Austria. *Remote Sens.* **2019**, *11*, 2018. [CrossRef]

24. Fang, J.; Yang, W.; Luan, Y.; Du, J.; Lin, A.; Zhao, L. Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. *Atmos. Res.* **2019**, *223*, 24–38. [CrossRef]

25. Sunilkumar, K.; Yatagai, A.; Masuda, M. Preliminary Evaluation of GPM-IMERG Rainfall Estimates Over Three Distinct Climate Zones With APHRODITE. *Earth Space Sci.* **2019**, *6*, 1321–1335. [CrossRef]

26. Mazzoglio, P.; Laio, F.; Balbo, S.; Boccardo, P.; Disabato, F. Improving an Extreme Rainfall Detection System with GPM IMERG data. *Remote Sens.* **2019**, *11*, 677. [CrossRef]

27. Tao, Y.; Duan, Q.; Ye, A.; Gong, W.; Di, Z.; Xiao, M.; Hsu, K. An evaluation of post-processed TIGGE multimodel ensemble precipitation forecast in the Huai river basin. *J. Hydrol.* **2014**, *519*, 2890–2905. [CrossRef]

28. Gebregiorgis, A.S.; Kirstetter, P.-E.; Hong, Y.E.; Gourley, J.J.; Huffman, G.J.; Petersen, W.A.; Xue, X.; Schwaller, M.R. To What Extent is the Day 1 GPM IMERG Satellite Precipitation Estimate Improved as Compared to TRMM TMPA-RT? *J. Geophys. Res. Atmos.* **2018**, *123*, 1694–1707. [CrossRef]

29. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Oxford Academic: New York, NY, USA, 2011; ISBN 9780123850225.

30. McGuffie, K.; Henderson-Sellers, A. *The Climate Modelling Primer*, 4th ed.; Wiley-Blackwell: Chichester, UK, 2014; ISBN 9781119943365.

31. Trenberth, K.E. Climate Diagnostics from Global Analyses: Conservation of Mass in ECMWF Analyses. *J. Clim.* **1991**, *4*, 707–722. [CrossRef]

32. Kidd, C.; Dawkins, E.; Huffman, G. Comparison of Precipitation Derived from the ECMWF Operational Forecast Model and Satellite Precipitation Datasets. *J. Hydrometeorol.* **2013**, *14*, 1463–1482. [CrossRef]

33. Wang, J.; Wang, H.-J.; Hong, Y. Comparison of satellite-estimated and model-forecasted rainfall data during a deadly debris-flow event in Zhouqu, Northwest China. *Atmos. Ocean. Sci. Lett.* **2016**, *9*, 139–145. [CrossRef]

Article

Correcting Position Error in Precipitation Data Using Image Morphing

**Camille Le Coz [1,*], Arnold Heemink [2], Martin Verlaan [2,3], Marie-claire ten Veldhuis [1]
and Nick van de Giesen [1]**

[1]   Faculty of Civil Engineering and Geosciences, Water Resources, Delft University of Technology,
     2628 CN Delft, The Netherlands; J.A.E.tenVeldhuis@tudelft.nl (M.-c.t.V.);
     N.C.vandeGiesen@tudelft.nl (N.v.d.G.)
[2]   Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft, 2628 CN Delft,
     The Netherlands; A.W.Heemink@tudelft.nl (A.H.); Martin.Verlaan@deltares.nl (M.V.)
[3]   Deltares, 2600 MH Delft, The Netherlands
*   Correspondence: c.m.l.lecoz-1@tudelft.nl

Abstract: Rainfall estimates based on satellite data are subject to errors in the position of the rainfall events in addition to errors in their intensity. This is especially true for localized rainfall events such as the convective rainstorms that occur during the monsoon season in sub-Saharan Africa. Many satellite-based estimates use gauge information for bias correction. However, bias adjustment methods do not correct the position errors explicitly. We propose to gauge-adjust satellite-based estimates with respect to the position using a morphing method. Image morphing transforms an image, in our case a rainfall field, into another one, by applying a spatial transformation. A benefit of this approach is that it can take both the position and the intensity of a rain event into account. Its potential is investigated with two case studies. In the first case, the rain events are synthetic, represented by elliptic shapes, while the second case uses real data from a rainfall event occurring during the monsoon season in southern Ghana. In the second case, the satellite-based estimate IMERG-Late (Integrated Multi-Satellite Retrievals for GPM) is adjusted to gauge data from the Trans-African Hydro-Meteorological Observatory (TAHMO) network. The results show that the position errors can be corrected, while preserving the higher spatial variability of the satellite-based estimate.

Keywords: precipitation estimation; satellite-based precipitation; gauge data; IMERG; TAHMO; morphing; field displacement

1. Introduction

Precipitation is an important variable in weather and climate research and many other applications. Precipitation data are needed as input for hydrological models, for flood and drought monitoring or for water management in agriculture or power generation. However, estimating precipitation accurately is difficult because of its high spatial and temporal variability. This is especially true for sub-Saharan Africa, where most of the rainfall is produced during the monsoon season by convective rainstorms, which are very localized [1,2].

Rain-gauges are the most direct way to measure precipitation. However, the gauge networks in Africa are not dense enough to derive high resolution precipitation estimates. Indeed, the rain-gauge distribution is sparse in many African regions and their number has been decreasing in recent decades [3]. During the same period, many efforts have been made to derive precipitation estimates from satellite data. Satellites do not measure precipitation directly but have the advantage of covering

large areas. This is especially interesting for Africa where gauge networks are sparse and there are also almost no radar observations available.

There is an increasing number of satellite-based rainfall products, providing rainfall estimates at different spatial and temporal resolutions. Most rainfall products use additional sources of data, such as gauge estimates, for bias correction. Bias correction methods focus on correcting the intensity. However, the intensity is not the only possible error in precipitation. Rainfall events are coherent moving systems and, in the case of convective rainstorms, they are also very localized. This can lead to errors in the estimation of the position and shape of the rain events beside the errors in their intensity. For some applications, such as hydrological modeling [4,5], flash flood warnings [6] or data assimilation in a numerical weather model [7,8], detecting the correct location of the rain events can be as important as their intensity.

The position errors in weather forecast models, including precipitation, have been taken into account in the field of forecast verification. Several spatial verification approaches have been developed [9,10]. They can be divided into four categories: neighborhood, scale-decomposition (e.g., References [11–13]), object- (or feature-)based (e.g., References [14–16]) and field deformation. In this study, we focus on a method belonging to the latter category. We now give an overview of field deformation method used for weather-related variables. Field deformation methods are based on a spatial mapping or displacement that makes a field (e.g., forecast) more similar to a target field or observation. The deformation is determined by minimizing a cost function. The Feature Calibration and Alignment technique (FCA [17–19]) is one of these methods. FCA has also been used for correcting position errors in cloud or water vapor related fields in the framework of data assimilation. For instance, References [18,20] corrected position error in a numerical weather model background fields using integrated water vapor measurements from satellite. In Reference [21], the FCA is used as a prepossessing step of an ensemble-based variational assimilation scheme for (satellite) brightness temperature. Reference [22] tested this method with several types of observations—integrated water vapor, lower level pressure, brightness temperature and simulated radar reflection. Other feature alignment techniques have been developed and used in data assimilation schemes, such as Reference [23] (for simulated radar observation), References [24,25] (for some idealized cases). The FCA technique has been applied directly to rainfall data in Reference [19]. They corrected rainfall estimates derived from SSM/I data with ground-based radar estimates. They illustrated the performance of their approach for different types of rainfall events, such as Hurricane Andrew, a squall line in Oklahoma and coastal rainfall in Australia.

Some field deformation methods for spatial verification originate from image processing, such as the optical flow techniques developed in Reference [26,27] or in Reference [28] and evaluated in References [29,30]. Image warping has also been used in data assimilation frameworks. Reference [31] assimilated integrated water vapor from satellite to improve a numerical weather model forecast. However, this method requires the manual selection of pairs of points to perform the image warping. Reference [32] combined image morphing with an ensemble Kalman filter for a wild fire model. They use an automatic registration technique that only requires two fields to derive the displacement field, without any manual specification needed. Using the same morphing and registration method, a morphing fast Fourier transform (FFT) EnKF for radar precipitation is described in Reference [33]. However, this morphing FFT EnKF is not implemented and applied to rainfall data.

This present study investigates the use of the morphing approach for the position correction of rainfall estimates, using the approach proposed by References [32,33]. While the goal of Reference [33] was to derive a method to assimilate radar precipitation into a numerical weather model, we aim to correct the position error of satellite-based precipitation estimates using gauge measurements. We apply the morphing approach to real precipitation data, namely the (non-gauge adjusted) IMERG-Late estimates and the new Trans-African Hydro-Meteorological Observatory (TAHMO) gauge network.

The morphing and automatic registration methods, including the case of irregularly spaced observations, are described in Section 2. The morphing approach is applied to two cases. The first case uses synthetic rainfall events represented by ellipses (Section 3.1). The second case is a real rainfall event occurring in southern Ghana during the monsoon season (Section 3.2). Both the convergence of the automatic registration and the performance of the warping are examined in Section 4. The results of the two cases are compared and discussed in Section 5, before the conclusion in Section 6.

2. Methodology

In this section, we define the image registration and morphing processes, before focusing on the implementation of an automatic registration procedure. We use the framework described by Reference [33].

2.1. Definitions

Let u and v be two signals (or images) defined on a domain $D \subset \mathbb{R}^2$ and $T : (x, y) \in \mathbb{R}^2 \mapsto \left(T_x(x, y), T_y(x, y)\right) \in \mathbb{R}^2$ be a mapping function. The goal of **image registration** is to determine a spatial mapping T such that, $\forall (x, y) \in D$,

$$
\begin{aligned}
v(x, y) &\approx u \circ (I + T)(x, y) \\
&\approx u\left[(I + T)(x, y)\right] \\
&\approx u\left[x + T_x(x, y), y + T_y(x, y)\right]
\end{aligned}
\tag{1}
$$

where I is the identity function.

There can be several mappings T that meet the requirement $v \approx u \circ (I + T)$. Especially in areas without rainfall, the mapping T is not unique. We define three criteria to characterize one optimal mapping:

$$
T \approx 0 \tag{2}
$$
$$
\nabla T \approx 0 \tag{3}
$$
$$
\nabla \cdot T \approx 0 \tag{4}
$$

That is, the optimal mapping has to be as small, smooth and divergent-free (i.e., it is not shrinking or expanding the field) as possible.

Several approaches have been used to define the optimality of the mapping. For the FCA method applied to precipitation, Reference [19] use smoothness and barrier conditions. Contrary to our condition on the magnitude (Equation (2)), their barrier does not impact small scale displacements. Using the FCA for data assimilation, References [20,22] added two more constraints, one on the magnitude and one on the divergent. Reference [25] did not use any magnitude or barrier approach and only had constraints on the gradient and the divergence. Our constraints on the magnitude and on the smoothness are the same as those used in Reference [32]. Constraints on the divergence were used in several similar field distortion methods [20,22,25]. Thus, we also added a third constraint on the divergence in order to observe its impact. A short sensitivity study on the impact of these three coefficient is presented in Appendix B.

Image warping is the distortion of an image based on a spatial transformation of the domain. Warping can be used to transform an image into another one by using the spatial mapping T obtained from the registration method. The mapping T is gradually applied to the original image u as follows:

$$
u_{\text{warp}(\lambda)} = u \circ (I + \lambda T) \qquad 0 \leq \lambda \leq 1 \tag{5}
$$

Warping works well when the residual $v - u \circ (I + T)$ is small, which is not the case when the images u and v have different intensities for example. It is a spatial transformation. It only acts on the

coordinates, it does not modify the intensity of the image u. On the other hand, **Cross-dissolving** only acts on the intensity. It fades two images u and v into each other:

$$u_{\text{diss}(\lambda)} \quad = \quad u + \lambda(v - u) \quad 0 \le \lambda \le 1 \tag{6}$$

Image morphing combines warping and cross-dissolving to account for both the spatial distortion and the difference in intensity:

$$u_{\text{morph}(\lambda)} \quad = \quad (u + \lambda r) \circ (I + \lambda T) \quad 0 \le \lambda \le 1 \tag{7}$$

where r is the residual:

$$r \quad = \quad v \circ (I + \lambda T)^{-1} - u \quad 0 \le \lambda \le 1 \tag{8}$$

With this formula of u_λ, we obtain $u_{\text{morph}(0)} = u$ and $u_{\text{morph}(1)} = v$.

2.2. Automatic Registration

The spatial mapping T used for the image morphing is determined by the image registration. Several registration methods are available. However, many of them require to define manually a set of corresponding points from the images u and v. We are interested in an automatic registration procedure that only needs the images u and v as inputs without any extra specifications. This requires the images to be similar enough for the automatic registration procedure to work.

We use the method described by Reference [33] based on the minimization of a cost function J with respect to the mapping T. The cost function can be divided in two terms (Equation (9)). The first one (J_o) represents the mapping error between the displaced original signal $u \circ (I + T)$ and the target signal v. The second one (J_b) is a background term that consists of the three criteria for 'optimal' mapping given in Equations (2)–(4). These three criteria are used as weak constraints.

$$
\begin{aligned}
J(T) &= J_o(T) + J_b(T) \\
J_o(T) &= \|v - u \circ (I + T)\| \\
J_b(T) &= C_1\|T\| + C_2\|\nabla T\| + C_3\|\nabla \cdot T\|
\end{aligned}
\tag{9}
$$

where C_1, C_2 and C_3 are three coefficients determined empirically and $\|.\|$ is the L^2-norm.

The minimization problem is solved iteratively, for T defined on increasingly fine grids. The iterative approach has two advantages. It helps reduce the computational cost and avoids the local minima problem (see below).

In our application, the domain D is rectangular. It can be represented by different uniform grids. The regular $n_x \times n_y = n$ grid on which u, v and $u_{\text{morph}(\lambda)}$ are given is called the *pixel grid D_n*. The mapping function T is defined on a set of coarser grids D_i ($i = 1, ..., I$), called *morphing grids*. It is then represented by two gridded arrays (one for T_x and one for T_y). The grids D_i are uniform $(2^i + 1) \times (2^i + 1) = m_i$ grids (for $i = 1, ..., I$) covering the domain D. For $i = 1, ..., I$, the mapping T discretized on D_i is noted T_i.

The signals u and v, and so the observation term J_o of the cost function, are discretized on the pixel grid D_n. The background term J_b is discretized on the morphing grid D_i. We use the second order central scheme except at the boundaries where the first order backward or forward schemes are used. We use bilinear interpolation to estimate the value of u and v on the distorted grid (e.g., $u \circ (I + T)$) and to interpolate T on the different morphing grids D_i.

The finest morphing grid D_I does not need to be the same as the pixel grid D_n. On the contrary, it is computationally advantageous when the morphing grid D_I has a much coarser resolution. When the number of nodes m_i of the morphing grids is much smaller than the number of nodes n of the pixel grid,

solving the minimization problem on the set of morphing grids D_i is less computationally expensive than to solve it for T defined on the high resolution pixel grid D_n.

Algorithm

The algorithm iterates over the morphing grids D_i ($i = 1, ..., I$), starting on the coarsest 3-by3 grid D_1, until it reaches the finest morphing grid D_I. For each iteration, the three main steps are similar to those in Reference [32] and are illustrated in Figure 1.

1. **Smoothing** of the images u and v: the images are smoothed by convolution with a 2D-Gaussian

$$G_{2D}(x,y) \quad = \quad \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right),$$

where $\sigma = 0.05/(2^{2i} + 1)$. The finer the grid D_i is, the narrower the Gaussian is. Thus, for small i, the fine features are ignored and the focus is given to the large-scale ones. When i increases, more and more fine features are taken into account. This way, T_i for small i will make the larger features match. Then, for increasing i, more and more detailed images are matched.

The cost function J is often non-convex with respect to T and so can have several local minima. The smoothing combined with the hierarchy of grids reduce the local minima problem. They ensure that the large-scale features are fitting first, hence avoiding local minima.

After the smoothing, the two fields are normalized such that their maximum is the same. The images obtained after smoothing and normalization are noted as $\tilde{u}_i$ and $\tilde{v}_i$.

2. **Initialization**: solving the minimization problem on grid D_i requires a first guess T_i^{fg}. For $i = 1$, T_1^{fg} is set to zeros, that is, no deformation. For $i = 2, ..., I$, the mapping T_{i-1}^* obtained by solving the minimization problem on grid D_{i-1} is interpolated into the grid D_i and used as the first guess T_i^{fg}.

3. **Optimization**: The actual minimization problem to be solved is based on the smoothed fields, that is, $J_o(T) = \|\tilde{v}_i - \tilde{u}_i \circ (I + T)\|$. Contrary to Reference [32], we solved the minimization problem for all the nodes at the same time.

There is a number of inequality constraints on this minimization problem, due to our requirements of invertibility. An iterative barrier approach is used to transform this constrained minimization problem into an unconstrained one [34,35]. In the barrier approach, the minimization is applied to a penalized cost function $J_p(T) = J(T) + \beta \sum_h C_h(T)$, where C_h are the constraint functions and β the barrier coefficient (over which we iterate when the constraints are not respected). The constraints and the minimization method are described with more details in Appendix A.
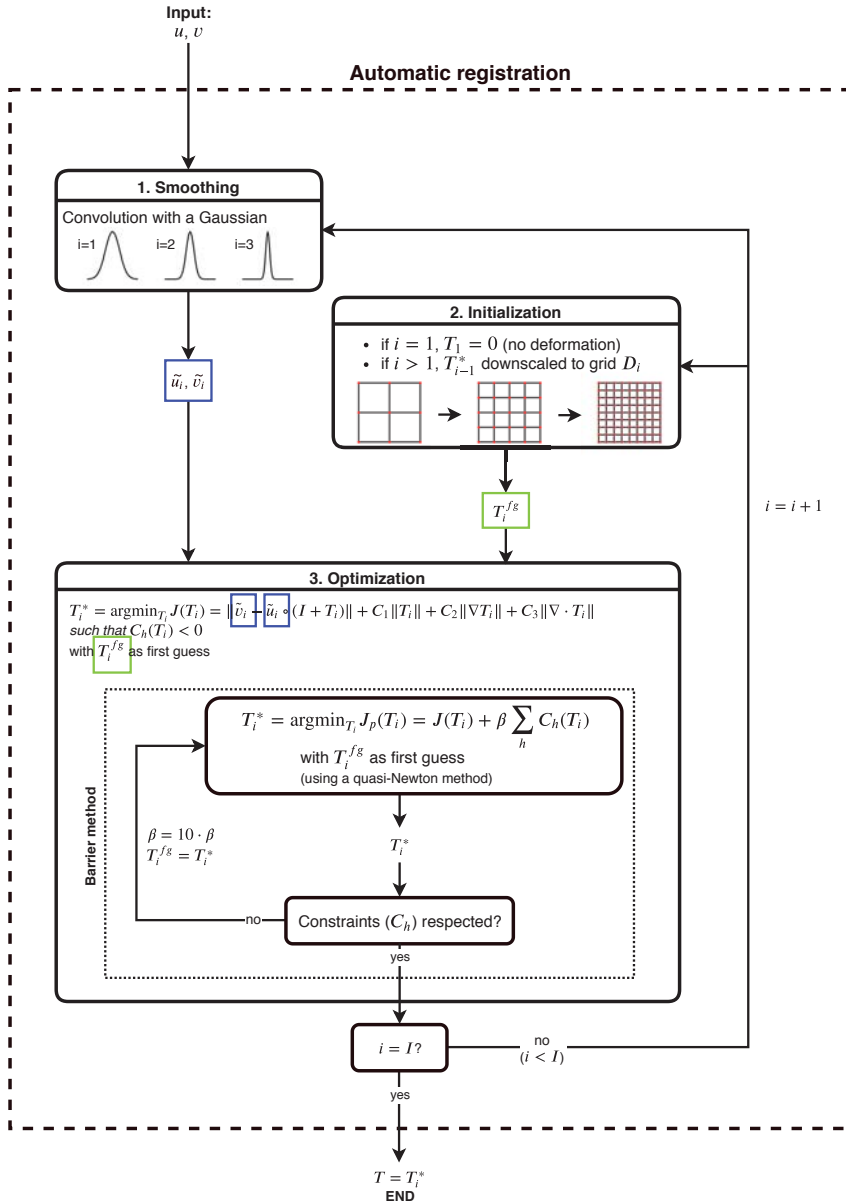
Figure 1. Algorithm for the automatic registration.

The python scripts for the automatic registration and the morphing are available online (Supplementary https://github.com/clecoz/precipitation-morphing.git). The scripts permit to reproduce the results for the synthetic case described below.

2.3. Dealing with Irregularly Spaced Observations

The automatic registration algorithm described above assumes that both signals u and v are on the same regular grid. However, in practice, one might deal with irregularly spaced observations, such as rain-gauge data.

In such a case, the observations are interpolated on the same regular grid, using kriging (details about the kriging are given in Section 3.2). In the remainder of the article, we will refer to the gauge interpolation as "kriging", while "interpolation" will refer to the bi-linear interpolation used in the automatic registration and morphing. The cost function J fro Equation (9) is modified to take the unequal coverage of the domain into account. A mask function M is added in the first term of J:

$$J(T) \quad = \quad \|M \cdot (v - u \circ (I + T))\| + C_1\|T\| + C_2\|\nabla T\| + C_3\|\nabla \cdot T\| \tag{10}$$

were $\cdot$ is the element-wise matrix multiplication. The mask function is defined such that it is equal to 1 in a given perimeter around the observations and zero everywhere else. So, the difference $v - u \circ (I + T)$ for the grid points far from any observation does not weigh in the cost function J.

3. Study Cases

3.1. Synthetic Case for Algorithm's Validation

A synthetic case is used to investigate the convergence of the automatic registration algorithm and to validate the morphing. The synthetic precipitation fields are generated on a regular $n_x \times n_y$ grid, with $n_x = n_y = 65 (= 2^6 + 1)$. The synthetic rainfall events are represented by ellipses and are added to a zero precipitation field (Figure 2). The fields u and v have two rainfall events each. However, they are at different positions. The distortion between the fields include both a rotation and a shear. The events also differ by their intensity and shape. The event in the lower right corner has the same shape in both fields u and v, but has a difference of intensity of 5 mm/h. The event in the upper-left corner has the same intensity in both u and v, but the ellipse has a different inclination.
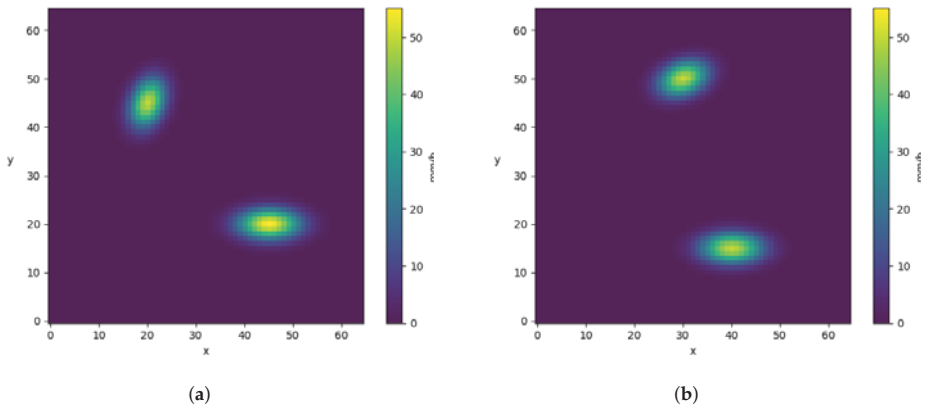

Figure 2. Precipitation fields u and v for the synthetic case, shown respectively in (**a**) and (**b**).

3.2. Southern Ghana Case

In this case, we apply the automatic registration and the warping to real precipitation data. Our goal is to gauge-adjust a satellite-based estimate with respect to the location of the rain events. We assume the gauge measurements to be more accurate, but the IMERG has a higher spatial variability that the gauges cannot reproduce because of their network's density. Thus, our goal is to keep the spatial variability of IMERG, while correcting the position mismatch with respect to the gauge. For this real case, we will only look at the warping for this case. The morphing would make the warped field similar to the gauges, even in areas without gauges.

The study area is a square domain over southern Ghana encompassing the Ghanaian cocoa region. This domain has been chosen because of the particular high density of the Trans-African Hydro-Meteorological Observatory (TAHMO) network in this area (Figure 3).

Southern Ghana has two rainy seasons. The main one extends from March to mid-July and the second one occurs during September and October. We chose a rainfall event during the main monsoon season. More precisely, we selected one hour during this event (from 18:00 to 19:00 of 22 April 2018). Given the spatial resolution of IMERG ($0.1°$ lat/lon), hourly accumulation is in good agreement with the rainfall spatial and temporal variability [5]. Longer time scale would make it more difficult to identify the individual events and very few precipitation datasets are available at a sub-daily scale.
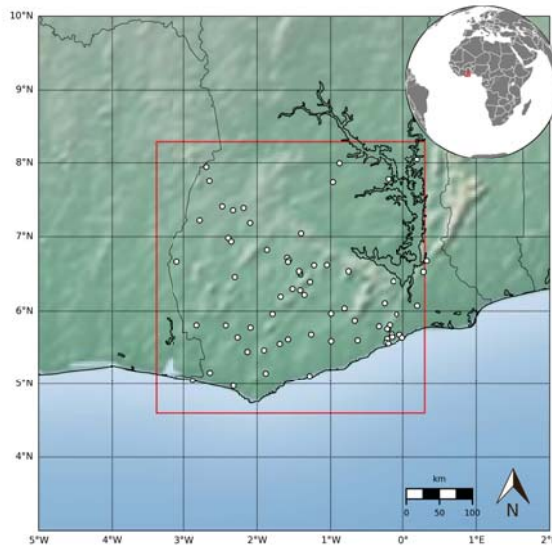
Figure 3. Study domain (red rectangle) and the TAHMOstations available within the domain (white dots).

3.2.1. Precipitation Datasets

IMERG (Integrated Multi-satellitE Retrievals for GPM) is a high resolution global precipitation product produced by NASA as part of the Global Precipitation Measurement (GPM) mission [36–38]. IMERG merges several satellite estimates from infrared, passive-microwave and satellite-radar. Three versions are available at half-hourly and $0.1°$ lat/lon resolution: the Early, Late and Final runs. The Final run is gauge-adjusted at monthly scale with the GPCC (Global Precipitation Climatology Centre) gauge product. This product is not available for the Early and Late runs, which have

respectively 5 h and 15 h latency. Instead, they are climatologically adjusted to the Final run. So, they indirectly incorporate past gauge data through this climatological adjustment, but are independent of recent rainfall measurements. The warping will be applied to IMERG-Late run, on a 37×37 grid points area corresponding to the study domain.

The Trans-African Hydro-Meteorological Observatory (TAHMO [39]) initiative aims to develop a dense network of 20,000 hydro-meteorological stations in sub-Saharan Africa (equivalent to one station each 30 km). These low-cost stations measure the standard meteorological variables, including precipitation, at a high temporal resolution (5 min). The current TAHMO network contains over 500 stations, mainly in West and East Africa. Data from 66 TAHMO stations are available within the study area for the selected hour.

IMERG-Late estimates and TAHMO measurements for the selected hour, i.e., for the 22 April 2018 between 18:00:00 and 19:00:00, can be seen in Figure 4. Both show one main rainfall event, with a maximum of 45 mm/h according to IMERG-Late. TAHMO recorded a slightly higher maximum of 53.45 mm/h at a gauge located South-West of IMERG's peak.
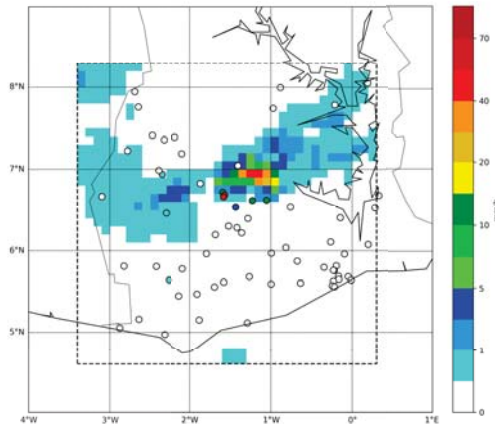
Figure 4. IMERG-Late (background) and TAHMO (circles) accumulated rainfall between 18:00:00 and 19:00:00, within the study domain (dotted line).

3.2.2. Data Pre-Processing

TAHMO measurements are given on an irregular grid, while the automatic registration algorithm assumes that all the data are defined on a regular grid. The TAHMO measurements were kriged on the same 0.1° lat/lon grid as IMERG estimates (Figure 5). We used ordinary kriging with a square root transform [40]. The ordinary kriging was done with *PyKrige*, a kriging toolkit for python (https://pypi.org/project/PyKrige/). We choose an exponential variogram with a sill of 1.0 (mm^2/h^2), a range of 1.5° and a nugget of 0.01 (mm^2/h^2).

Contrary to the previous cases, this case uses real noisy data. Hence, some pre-processing (beside the kriging) is necessary before applying the automatic registration. First, light precipitation (<0.1 mm/h) is removed from both fields. Lighter precipitation is in general more difficult to detect, hence subject to large uncertainties. Second, we add a zero-precipitation padding area around the study domain. This area is added to avoid problems near the boundary due to the minimization constraints. The extended domain corresponds to a $n_x \times n_y$ pixel grid, with $n_x = n_y = 65$. Finally, we normalized

the two fields such that the maximum rainfall is equal to 50 mm/h for both. Figure 6a,b show the fields u (IMERG-Late) and v (TAHMO) after the three pre-processing steps described above. The study domain is delimited by the dotted line, while the extended domain including the zero-padding area is delimited by the solid line. The extended domain corresponds to a regular $n_x \times n_y$ grid, with $n_x = n_y = 65$. These modified precipitation fields will be used as inputs for the automatic registration procedure to find the mapping T. Once the mapping T is found, it will be applied to the original fields u (Figure 4) on the extended domain.
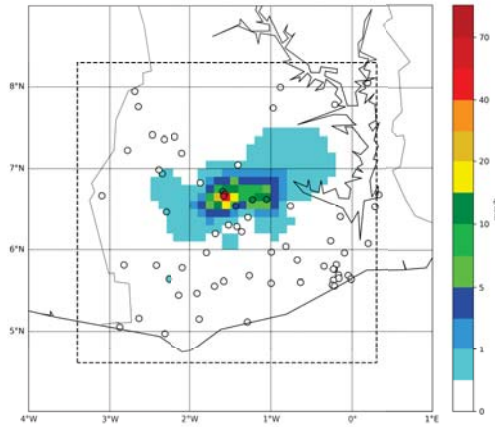
Figure 5. TAHMO measurements (circles) and TAHMO kriged (background) within the study domain (dotted line) for the selected event.

(a)   (b)
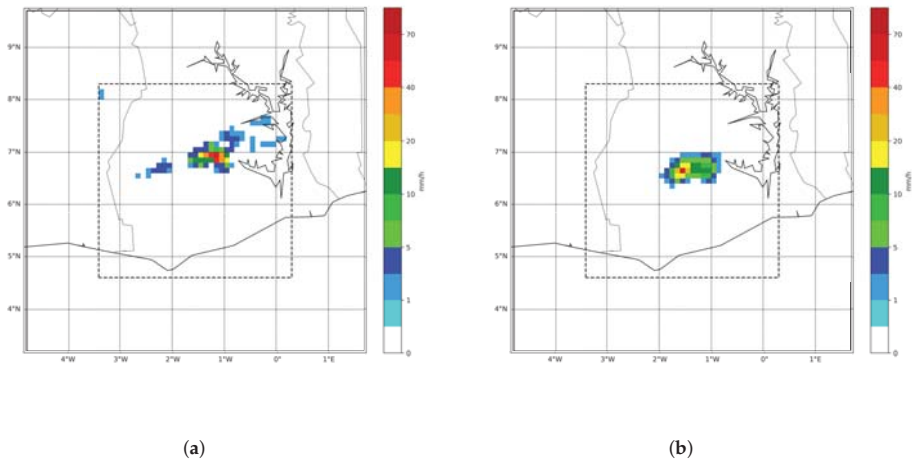
Figure 6. Fields u (IMERG-Late) and v (TAHMO) after pre-processing and used as inputs. The dotted line delimits the study domain and the solid line the extended one, with the zero-padding area in-between. (**a**) Field u (IMERG) after pre-processing. (**b**) Field v (TAHMO) after pre-processing.

4. Results

4.1. Synthetic Cases

The coefficients C_1, C_2 and C_3 in the cost function J are set empirically to $C_1 = 1$, $C_2 = C_3 = 10$. We will first look at the convergence of the minimization problem on each domain D_i for $i = 1, ..., I$, with $I = 5$ (so that $m_I = 33 \times 33 < n = n_x \times n_y = 65 \times 65$, see Section 3). Then, we will look at the performance of the automatic registration and morphing procedure.

The coefficients C_1, C_2 and C_3 influence the value of the cost function, and so the convergence. However, their impact on the mapped field is limited and mainly affect the area with no or low precipitation. For example, multiplying the coefficients by two will give similar results in terms of mean absolute error and root mean square error (see Appendix B).

4.1.1. Convergence

For each domain D_i ($i = 1, ..., 5$), we investigate the convergence of both the inner loop (from the barrier approach, that is, the iterations on the coefficient β) and of the outer loop (from the L-BFGS-B minimization method) for the synthetic case (Figure 2).

The value of the cost function before and after minimization for each step i, as well as the number of iterations needed by the L-BFGS-B method and the barrier approach to reach convergence, are shown in Table 1. The iterations on the coefficient β are performed only if some constraints are violated. For the ideal case described above, iterations over β are required for grids D_2 and D_5 (with a maximum of 3 iterations for D_3). The number of iterations needed by the L-BFGS-B method varies depending on the domain D_i and the coefficient β. It is interesting to observe that less iterations are necessary for increasing β. Domain D_3 necessitates the highest number of iterations and D_1 the lowest (1308 and 81 iterations respectively). These numbers are specific to this case and will vary depending on the input fields u and v, and on the coefficients C_1, C_2 and C_3. However, they show that convergence may require a high number of iterations and that this number is difficult to predict.

Table 1. Optimization results after each step $i = 0, ..., 5$ for the synthetic case. The number of iterations needed for the barrier approach (β iterations) and for the L-BFGS-B method are given separately. The total number of iterations correspond to the sum of the L-BFGS-B iterations for each β iteration. The cost function J_p is evaluated before and after optimization (i.e., for the first guess T_i^{fg} and the 'optimal' grid T_i^*). The latter is also separate into three terms, the mapping error (J_m), the background error (J_b) and the penalization term (not shown here because of its value close to zero).

i	1	2	3	4	5
Number of iterations					
- Total	81	996	1308	502	443
- β iter.	1	3	1	1	2
- L-BFGS-B iter. for each β	81, $\beta = 1$	956, $\beta = 1$	1308, $\beta = 1$	502, $\beta = 1$	263, $\beta = 1$
		36, $\beta = 10$			180, $\beta = 10$
		4, $\beta = 100$			
Cost function					
$J_p(T_i^{\text{fg}})$	213.169	54.722	40.568	24.278	16.371
$J_p(T_i^*)$	30.058	18.408	19.193	14.790	13.191
$J_o(T_i^*)$	27.552	16.352	17.847	13.399	11.734
$J_b(T_i^*)$	2.507	2.056	1.345	1.391	1.457

The reduction of the cost function is more important on the coarser domains. It is divided by 7 on domain D_1, by 3 on domain D_2 and by 2 on domain D_3 (Table 1). This is probably due to better first guesses, since the displaced grid obtained on a domain D_i is used to initialize the finer grid D_{i+1}. The observation error term J_o of the cost function is reduced greatly after optimization on the four coarsest domains (from 27.552 after optimization on domain D_1 to 13.399 on domain D_4). However, it increases from domain D_2 to D_3, before decreasing again on domain D_4. This increase could be due to the smoothing which is narrower for increasing i, making the position error more apparent in the mapping error. This could also be due to the background error term (J_b), which larger decrease is also from domain D_2 to D_3.

To investigate further the impact of each domain D_i on the automatic registration procedure, we examine the mean absolute error (MAE) of the morphed ($u_{\text{morph}} = u_{\text{morph}(\lambda=1)}$) and warped ($u_{\text{warp}} = u_{\text{warp}(\lambda=1)}$) signals for different values of I (shown in Table 2). The MAE of the morphed signal is divided by around 4.2 from $I = 1$ to $I = 2$ and by 2.6 from $I = 2$ to $I = 3$. Domains D_4 and D_5 have a smaller impact on the MAE (MAE divided by around 1.3). Similar trend can be observed for the MAE of the warped signal, but with a smaller decrease. The MAE is divided by 1.9 from $I = 1$ to $I = 2$ and from $I = 2$ to $I = 3$. Increasing I from 4 or 5 has only a minor effect on the MAE of the warped signal.

Table 2. MAE (in mm/h) of the morphed (u_{morph}) and warped (u_{warp}) signals obtained at different steps i, for the synthetic case.

i	1	2	3	4	5
$\text{MAE}(u_{\text{morph}} - v)$	0.678	0.161	0.062	0.050	0.036
$\text{MAE}(u_{\text{warp}} - v)$	0.531	0.281	0.147	0.137	0.133

The number of variables to optimize on a domain D_i increases with increasing i. So, solving the minimization problem becomes more computationally expensive. Moreover, we saw that domain D_5 has a smaller impact on the cost function and on the MAE of the morphed signals. Thus, in order to limit the computational cost, the finer morphing grid will be set to $I = 4$ in the following validation section.

4.1.2. Validation

In this subsection, we are looking at the performance of the automatic registration and of the morphing procedure on the synthetic case for $I = 4$.

From the automatic registration procedure, we obtain a mapping function T that will be used to distort the image in the morphing process. The mapping T is shown in Figure 7a. The corresponding grid distortion is illustrated in Figure 7b which shows the regular $n_x \times n_y$ grid (on which u and v are defined) before and after the mapping has been applied. By comparing with the signals u and v (in Figure 2), one can see that the mapping was able to take into account both the rotation and the shear between the events. The biggest distortions occur around the two rainy peaks, while the distorted grid is more regular and closer to the original grid in the areas further away from the peaks. This is due to the first regulation term of the cost functions ensuring that the transformation T is as "small" as possible. The second regulation term is responsible for the smooth transition between these areas. The impact of the third regulation term is more difficult to see directly. The third term penalizes grid cell to shrinkage or expansion. Due to the locations of the rainfall events in the signals u and v, almost the entire domain is influenced by the mapping T. Large distortions near the boundary can raise some inconsistencies in the distorted grid, such as in the one in the lower-right corner in Figure 7b. A way to avoid such discrepancies near the boundary would be to add a padding area around the domain filled with zero precipitation.

(a)   (b)
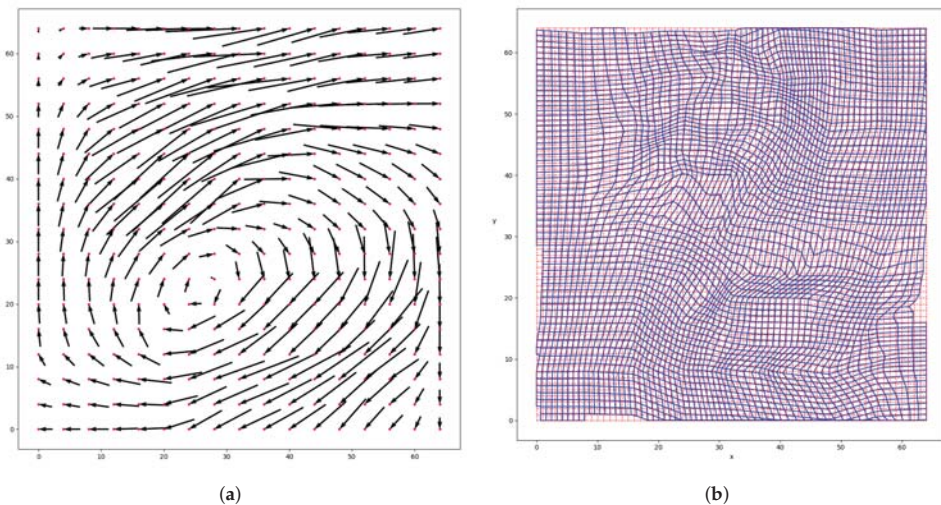
Figure 7. (**a**) Mapping T obtained from the automatic registration (with $C_1 = 0.1$ and $C_2 = C_3 = 1$) and (**b**) its effect on the pixel grid, for the synthetic case. (**a**) Mapping T. (**b**) Pixel grid D_n before (in red) and after (in blue) distortion by the mapping T.

Once the mapping T is obtained, it can be applied to the signal u. One can see the impact of the mapping by comparing the morphed signal u_{morph} (Figure 8b) to the original signal u (Figure 8a). The error between the morphed signal and the signal v is shown in Figure 8c. Figure 8c shows that the absolute error is small compared to the intensity of the rain events. This is also reflected by the low MAE in Table 1. Figure 8d shows that the errors have a clear delimitation, corresponding to the rain events of the target field v. For this synthetic case, the center of the two events in u_{morph} and v are corresponding exactly with respect of the position.

It is interesting to note that, while the lower event had an error on the intensity and an error on its location, the error of the morphed signal is lower around its peak than for the second (upper) events. The error of the warped field is close to 5 mm/h near this peak (not shown here). This shows

the advantage of morphing over warping when there is an intensity error. The upper event does not have any intensity error, but it has a shape error (i.e., the ellipses have different inclinations). The largest error in the morphed field u_{morph} occurs around its peak. The top is underestimated, while the surroundings are overestimated. This pattern would suggest an error due to the linear interpolation used for the morphing. This underestimation is greatly reduced when using a 2d spline instead of the current bi-linear interpolation (not shown here).
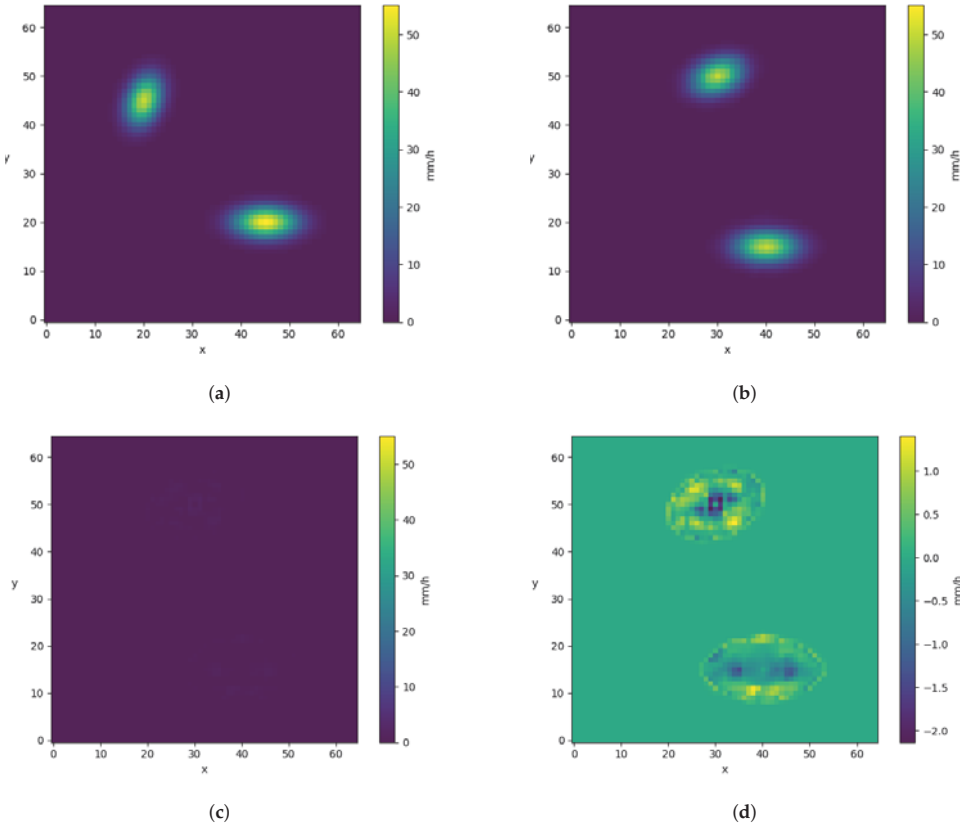
Figure 8. Morphing results: (**a**) original signal u, (**b**) morphed signal u_{morph}, (**c**) absolute error between the morphed signal and the target signal v and (**d**) error between the same two signals (using a different scale). (**a**) u. (**b**) u_{morph}. (**c**) $|u_{morph} - v|$. (**d**) $u_{morph} - v$.

4.2. Southern Ghana Case

The results described in this section are obtained for $I = 4$, with the regulation coefficients empirically set to $C_1 = 0.1$ and $C_2 = C_3 = 1$ (these coefficients are the same as for the synthetic cases). We define the mask function M such that $M = 1$ at the grid points where the kriging variance $\sigma_{kriging} < 0.5 \cdot \text{sill} = 0.5 \text{ mm}^2/\text{h}^2$ and $M = 0$ otherwise. The padding area is thus masked too.

4.2.1. Convergence

We first look at the convergence of the automatic registration procedure when applied to this real (noisy) case. Table 3 shows the number of iterations (nit) needed by the L-BFGS-B method to converge for each β, as well as the cost function before and after the minimization. Domain D_4 (for $\beta = 1$) necessitates the highest number of iterations and D_1 the lowest (1463 and 10, respectively). Iterations

on the coefficient β are preformed only on domain D_4. Contrary to the synthetic case, the reduction of the cost function is important on the finer morphing grid. It is divided by two on D_2 and D_3 and by five on D_4. In this case, the background term J_b is higher than the mapping error term J_m on all domains D_i. The cost function after optimization $J_p(T_i^*)$ increases from domain D_1 to D_4, mostly due to the background term J_b. The important increase of the mapping error $J_m(T_i^*)$ on domain D_4 is mostly due to the smoothing that reveals sharper features.

Table 3. Optimization results after each step $i = 0, ..., 4$ for the southern Ghana case. The number of iterations needed for the barrier approach (β iterations) and for the L-BFGS-B method are given separately. The total number of iterations correspond to the sum of the L-BFGS-B iterations for each β iterations. The cost function J_p is evaluated before and after optimization (i.e., for the first guess $T_i^{\mathbf{fg}}$ and the 'optimal' grid T_i^*). The latter has also been separated into three terms, the mapping error (J_m), the background error (J_b) and the penalization term (not shown here because of its value close to zero).

| i	1	2	3	4
Number of iterations				
- Total	10	60	433	1463
- β iter.	1	1	1	5
- L-BFGS-B iter. for each β	10, $\beta = 1$	60, $\beta = 1$	433, $\beta = 1$	1430, $\beta = 1$
				16, $\beta = 10$
				9, $\beta = 100$
				4, $\beta = 1000$
				4, $\beta = 10{,}000$
Cost function				
$J_p(T_i^{\mathbf{fg}})$	0.0516	0.0938	0.1068	0.3602
$J_p(T_i^*)$	0.0421	0.0424	0.0483	0.0679
$J_m(T_i^*)$	0.0168	0.0058	0.0060	0.0252
$J_b(T_i^*)$	0.0253	0.0376	0.0422	0.0427

4.2.2. Validation

The mapping T obtained from the automatic registration and its effect on the pixel grid D_n is shown in Figure 9. The field is shifted toward the South-West. The deformations are more important in the center and in the South of the domain, while being very small or null near the boundary. The area near the boundary corresponds to the padding area which is filled with zeros. The regulation terms are thus dominating the cost function in this area, especially the first one that ensure the mapping to be as small as possible. In contrast, the first term of the cost function is dominant in the center of the domain where the rainfall event is located. The second and third regulation terms ensure that the transition between these two areas is smooth.

This mapping T is then applied to the field u, the satellite estimate from IMERG, to correct the location of the rain event. Figure 10 shows the warped field u_{warp} and the TAHMO measurements. One can see that the location of the rainfall event is corresponding more to the gauge data than before the warping (Figure 4). We define the center of the event by the grid cell with the maximum precipitation. Using the kriged gauge field v as the truth, we compute the position error of the event's center before and after the warping. It decreases from 55,365 km to 22,096 km. This remaining error of 22,096 km is due to an error in the longitude. Indeed, the maximum rainfall is at the correct latitude, but is two grid cells to the East of the actual peak. This error in longitude can be explained by the internal structure of the event. By comparing the pre-processed fields in Figure 6, one can see that the peak is located in the eastern part of the event for IMERG (u), but in the West for the kriged gauge (v). This difference could be due to the kriging and gauge network density (no gauge measurement is available near the peak of the warped field).

(a)   (b)
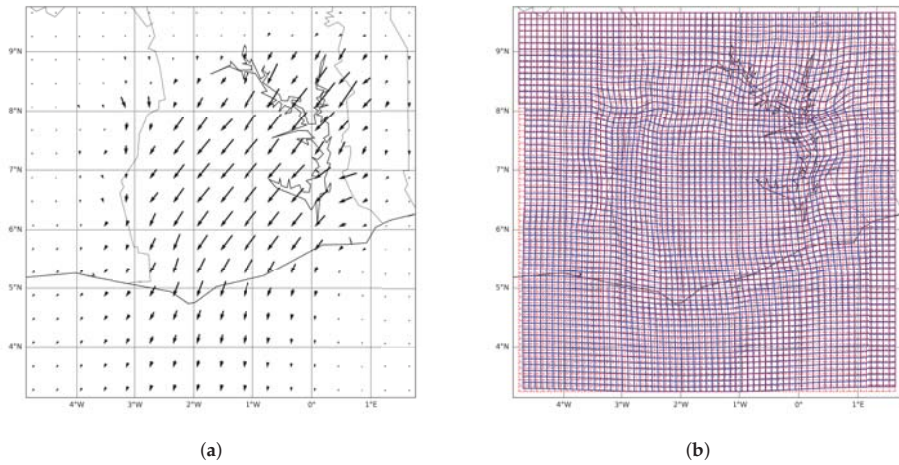
Figure 9. (**a**) Mapping T obtained from the automatic registration (with $C_1 = 0.1$ and $C_2 = C_3 = 1$) and (**b**) its effect on the pixel grid, for the southern Ghana case. (**a**) Mapping T. (**b**) Pixel grid D_n before (in red) and after (in blue) distortion by the mapping T.
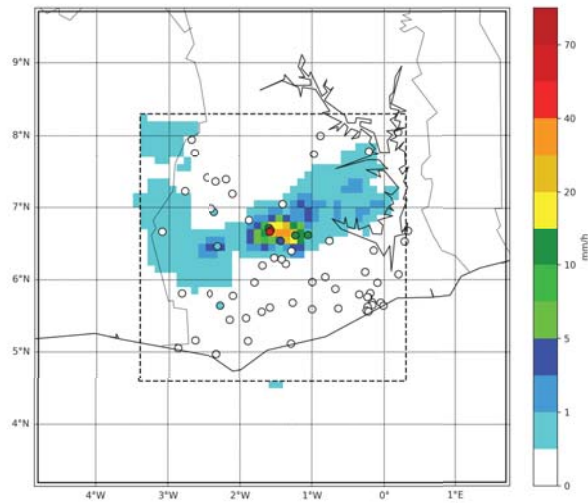
Figure 10. Warped signal u_{warp} (background) and TAHMO measurements (circles).

To investigate further the automatic registration, we compute the MAE and RMSE between the warped field and the target one, using the mappings T_i obtained at each step i (Table 4). Applying the mapping T_1, defined on the coarsest morphing grid, already greatly reduces the MAE and the RMSE compared to the original errors. Increasing the resolution of the morphing grid, further decreases the MAE and RMSE, except for the RMSE for $i = 2$. While the cost function is divided by five on domain D_4 (see Table 3), the reduction of the error from D_3 to D_4 is much smaller.

The target field v has been obtained by kriging the gauge data. Hence, some interpolation errors were introduced, especially in the areas far from the gauges. Comparison to the whole field v is not

representative, since it contains some large uncertainties. In the second part of Table 4, the warped field u_{warp} is estimated at the station locations (shown in Figure 11 for T_4) and is directly compared to the gauge measurements. As previously, a large part of error reduction is occurring on domain D_1. In total (i.e., for $i = I = 4$), the RMSE has been divided by almost two, and the MAE by 1.5.

Table 4. MAE and RMSE of the warped (u_{warp}) signal compared to the kriged TAHMO field (v) and to the gauge measurements, obtained at different steps i, for the southern Ghana case.

i	Before	1	2	3	4
Kriged field $\left(u_{\text{warp}} - v\right)$					
- MAE (mm/h)	0.1748	0.0968	0.0936	0.0886	0.0865
- RMSE (mm/h)	1.6476	0.9171	0.9496	0.8976	0.8565
Gauge measurements					
- MAE (mm/h)	2.0687	1.4365	1.4434	1.3849	1.3619
- RMSE (mm/h)	8.7053	5.5057	5.4984	4.9349	4.6690

Figure 11 shows the warped field at the station locations. The stations at which the gauge data and the warped field disagree are numbered. At the other stations, both TAHMO and the warped field estimate zero precipitation. The precipitation amounts at the 18 numbered stations according to TAHMO, IMERG-Late (u), IMERG-Final and the warped field (u_{warp}) are shown in Figure 12. They are ordered in decreasing order of precipitation with respect to TAHMO measurements. For stations 1 and 2 (with highest rainfall intensity), the warped signal is much closer to the measurements than IMERG-Late and IMERG-Final. However, u_{warp} underestimates the intensity and estimates more rainfall at station 2 than 1 (the opposite of TAHMO). This underestimation can be explained by two factors. First, the maximum of IMERG-Late u on our domain (=45 mm/h) is lower than the one recorded by TAHMO. Second, we used linear interpolation for the warping. The rainfall at station 3 is underestimated by the three satellite estimates. At the stations 4 and 5, both IMERG-Late and IMERG-Final underestimate while u_{warp} overestimates. IMERG misses the precipitation at station 6, while u_{warp} overestimates by almost a factor four. The remainder of the stations has very low or no precipitation according to TAHMO. The three satellite estimates are similarly low (less than 1 mm/h) at these stations, with a few exceptions: IMERG-Late and IMERG-Final at station 14 and u_{warp} at stations 8 and 12. The overestimation by u_{warp} at station 8 corresponds to a second lower peak present in the original field u but not in the target field v.
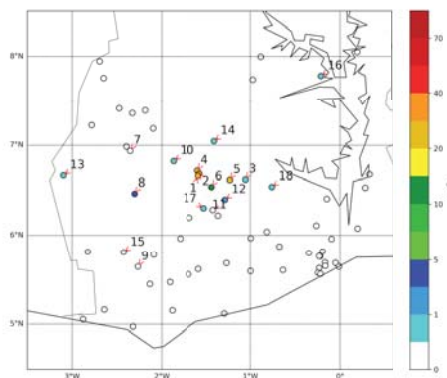
Figure 11. Warped signal u_{warp} at the stations location. The stations that differ from the TAHMO measurements are numbered.
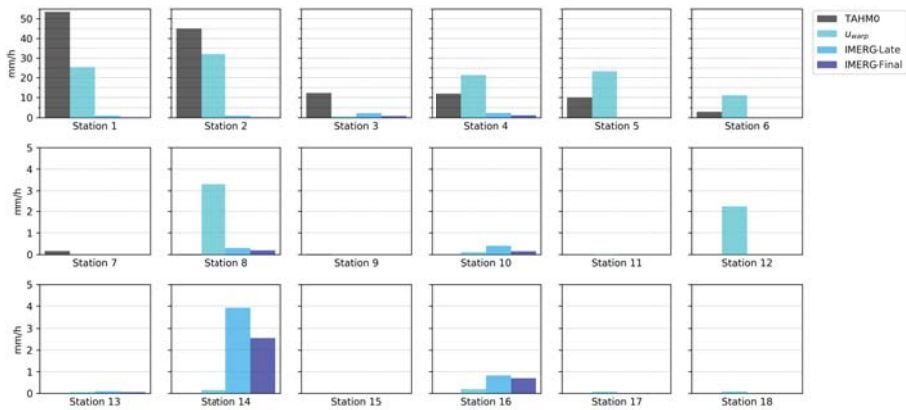
Figure 12. Rainfall (in mm/h) at 18 TAHMO stations according to TAHMO, IMERG-Late, IMERG-Final and the warped signal (u_{warp}).

5. Discussion

5.1. Convergence

The automatic registration algorithm converges for both the synthetic and the real (Southern Ghana) case. However, some differences can be noticed when comparing the optimization results shown in Tables 1 and 3.

The minimization method (L-BFGS-B) needs more iterations for the synthetic than for the real case for all steps i except for $i = 4$. Iterations on the coefficient β were needed for step $i = 2$ and $i = 5$ for the synthetic case, while the real case needed it for its finer step $i = 4$. The real case is noisier than the synthetic one, but the displacement between the field u and v is more straightforward. Indeed, for the synthetic case, the mapping T has to describe both a rotation and a shear. For the real case, the mapping T only has to represent a translation. On the coarser grids, the translation does not violate the constraints on the barrier. On the other hand, a rotation is more likely to violate the constraints and so to require iterations on β on coarser grids too. This shows that the number of iterations of both the minimization and barrier method depends on the input fields u and v, especially on the mapping complexity. It also depends on the coefficients C_1, C_2 and C_3, but the influence of the chosen coefficient values is limited (results not shown here).

For the synthetic case, the decrease in the cost function is more important for the first steps. This can be explained by better first guesses for the finer grids. On the contrary, for the real case, the decrease is more important in the last steps. While the events were identical in the synthetic case, they have different shapes in the real one. On the coarser grid, these differences are masked by the strong smoothing. They become more visible on the finer grid on which there is less smoothing. The sharper features being more sensitive to small position errors results in a higher cost function. The finer morphing grids allow the mapping to take the shape difference into account, on top of the position error. The reduction of the cost function is thus becoming more important.

The intermediate mappings T_i give information about the impact of the steps $i = 1, ..., I$. They were evaluated by looking at the MAE of the warped fields u_{warp} (Tables 2 and 4). In the real case, most of the MAE decrease is reached after the first iteration (divided by around 2). The decrease after the subsequent iterations is more marginal. The mapping on the coarser morphing grid D_1 is already able to capture reasonably well the displacement, that is, the translation toward the South-West. Thus, the finer morphing grids induce less improvement. In the synthetic case, the MAE is improved greatly after the first iteration too (divided by 7). However, the subsequent iterations continue to decrease the MAE (divided by 3 after step 2 and by 2 after step 3). The first morphing grids are too coarse to describe

accurately the complex displacement, which combine a rotation and a shear. Hence, increasing the resolution of the morphing grids improves the mapping T_i and thus the MAE. The more complex the displacement is, the finer the morphing grid needs to be and so the higher I has to be (i.e., the more steps i we need).

The computational time increases exponentially with the number of steps i. All computations shown here were done on a personal computer. For the synthetic case, the first four steps (i.e., $i = 1, 2, 3$ and 4) were completed in approximately 2 min, while the fifth iteration ($i = 5$) needed between 4 to 10 min (depending on the computer computational capacity). The real case requires fewer iterations than the synthetic one and so a shorter computational time ($\sim$1 min for $I = 4$ and $\sim$4 min for $I = 5$).

5.2. Validation

The automatic registration procedure converged, for both the synthetic and the real cases. It has been shown that the error between the two original fields was considerably reduced by applying the mapping T, even without bias adjustment (i.e., the warped field u_{warp}). An issue encountered in the synthetic case was the grid distortion near the domain boundary (Figure 7). Some inconsistencies can appear when the rainfall events are close to the boundary. This was solved in the real case by adding a padding area, filled with zero precipitation, around the domain. This padding area enables the mapping to have a smooth transition from the largest displacement near the events to (almost) none near the new extended boundary (Figure 9a).

The automatic registration produced reasonable coordinate mapping in these two cases. However, problems can arise if the dissimilarity between the two original fields are too strong. We do not have a method to quantify this problem beforehand. However, there are some minimum conditions, such as having the same number of events in both fields or the proximity of these events. The smoothing steps of the registration algorithm can also be increased or decreased to allow the events to move further or not. In this study, we did not push to the cases to the extreme to determine a feasibility threshold. The goal of this article was to prove the applicability of registration and morphing to precipitation data. A next step would be to apply it to other cases, including different rainfall regimes.

The warping succeeded in correcting the general position error between the fields u and v in both cases. In the real case, the shape of the event is different in the original fields. One can notice that the shape of the event is slightly altered by the mapping, but the internal structure stayed similar. The peak (rainfall above 20 mm/h) is larger in the field u than in the target field v, but the event (rainfall above 1 mm/h) has a larger longitudinal spread according to v (see Figures 4 and 5). This can explain some of the intensity differences at the station location (Figure 12). The stations 4 to 6 are in the center of the event, the station 4 is especially very close to the peak (stations 1 and 2). The overestimation by the warped field is related to the larger peak in the field u. Similarly, the underestimation at station 3 and the overestimation at station 12 is due to their location near the edge of the event. Stations 3 and 12 are located close to the East edge and South edge of the event respectively and so are affected by the spread difference of the event according to u and v.

The morphing has been evaluated only for the synthetic case. Table 2 shows the added advantages of morphing over warping. The MAE of the warped signal is larger by factor 2 on domain D_3 and by a factor of 2.7 on domain D_4 compared to the MAE of the morphed signal. This important difference is due to the intensity difference between the lower event in u and the one in v. However, this advantage decreases when the intensity difference decreases. When the intensity is the same for the two fields, warping is more advantageous. Without difference in intensity, the residual r in the morphing formula Equation (7) is unnecessary and only add numeric errors (because of the inverse transform $(I + T)^{-1}$ and the extra linear interpolation). The morphing was not tested for the real case because of the irregular nature of the observations. The uncertainty of the kriged field is high in large part of the domain where no gauges are available. We made assumptions on the spatial mapping through the three criteria for optimality. This allowed us to correct the position through the entire domain. However, we can not make similar assumptions for the intensity.

5.3. Applications

In this paper, we corrected a satellite-based precipitation estimate based on gauge measurements. This position correction could be particularly useful to pre-process satellite rainfall data for applications needing accurate rain event positioning. Image morphing can take both the position and the intensity into account but we do not recommend to correct both at the same time. The morphed estimate would then be comparable to the kriged gauge field, without any advantage of the more detailed spatial structure of the satellite observations. A two-step approach is preferred, with first the position correction using the warping and then a bias correction such as the additive-multiplicative one used by IMERG-Final. Such position correction could be particularly beneficial as a pre-processing step for hydrological modelling applications. Rainfall data is an important input for hydrological models and can have a large impact on their accuracy [41,42]. The correct positioning of rainfall events can be as crucial as their intensities, especially for the localized events.

The morphing can also be applied to rainfall fields from other sources, such as a numerical model. It can then be used for data assimilation. Two approaches are possible. The position correction can be applied as a first step before the usual data assimilation on the intensity [22,24]. It is also possible to assimilate both intensity and distortion at the same time, represented respectively by the residual r and the mapping T [32,33]. In the second case one can take full advantage of the morphing formulation. In this paper, we did not perform data assimilation as it was described in References [32,33]. Instead, we used a similar method to theirs to correct the position in a satellite-based estimate using gauge data. There are three main differences between our method and the morphing described in References [32,33]. First, they used two penalty terms to ensure the smoothness of the displacement field (based on its magnitude and gradient), we add a third penalty term based on the divergence. Second, they solved the minimization problem for one grid point at a time (i.e., they have several 1D minimization problems), while we solve it for all the grid points together (i.e., we have one multi-variable minimization problem). Finally, we extend the method to non-gridded observations. Contrary to radar data, the gauge measurements used in our study case are irregularly spaced (i.e., non-gridded). In Reference [33], the framework for assimilating radar rainfall using morphing is described but is not actually applied to real rainfall data.

The main limitation of image morphing is in fact the limitations of the automatic registration. As discussed above, it can fail if the fields are too dissimilar. It is also influenced by the three regulation coefficients C_1, C_2 and C_3. For example, in the case of a low intensity event, the regulation terms in J_b can dominate the cost function, not allowing the rain event to move. In this paper, we explore the feasibility of image morphing for position correction in precipitation estimates. However, we have not pushed to the extreme the cases to quantify its limits. This paper is meant as a proof-of-concept. The next step will be to extend the study to other cases, involving different rainfall regimes. Extreme cases should be included to determine the boundary within which the automatic registration succeeds

6. Conclusions

We have investigated the use of a morphing approach for the gauge-adjustment of satellite-based rainfall estimates with respect to position error. The morphing method, adapted from Reference [32], has been applied to two cases. Synthetic rainfall events, represented by ellipses, have been used to test the automatic registration and the morphing method. The second case, a convective rainfall event in southern Ghana, showed the potential of the method when applied to real, noisy precipitation data. We applied the position correction such that the gauge data were downscaled while keeping the high spatial variability of the satellite-based product. The rain events estimated by IMERG-Late were spatially shifted to match the gauge data. The morphing method can take both the intensity and the position of the rain events into account. This is an advantage compared to the traditional gauge-adjustment methods that are only looking at the intensity bias.

The automatic registration is able to represent different types of distortions. However, its performance of the registration depends on the degree of difference between the fields u and v and on the regulation coefficients. The more complex the distortion between the fields is, the more computationally expensive the registration is. For example, in the case of a simple distortion (such as a translation), it is possible to choose a smaller number of steps I. The minimization method at each step would also need fewer iterations. On the other hand, if the fields are too dissimilar, the registration can fail. The regulation coefficients also influence both the convergence and the result of the registration.

This paper explores the use of an image morphing method to correct location errors in precipitation estimates. The next step will be to extend the study to other case studies, including different rainfall regimes. It should also be pushed to more extreme cases to determine the method's limitations more precisely. For example, the regulation coefficients have been chosen empirically in this study. A next step will be to develop a more robust way to select them, for example by defining adaptive coefficients.

Appendix A. Optimization Step of the Automatic Registration Algorithm

The automatic registration is based on the minimization of the cost function J given in Equation (9). This is done in the third step of the algorithm (Figure 1). In this appendix, we give more information about this important step of the algorithm, with a focus on the constraints (Appendix A.1) and on the minimization method (Appendix A.2).

Appendix A.1. Constraints

This minimization problem include a certain number of constraints. These constraints come from two requirements. First, we require that the nodes stay within the domain D (i.e., $(I + T)(x_k, y_j) \in D \;\; \forall (x_k, y_j) \in D_i$. That is, the nodes on the boundary of the domain are allowed to move inside the domain but no nodes are allowed to leave it. These constraints can be seen as inequality constraints or as bounds. Second, the mapping $I + T$ has to be invertible. Each node (x_k, y_j) of the grid D_i is constrained to the domain between the points (x_{k+1}, y_j), (x_k, y_{j-1}), (x_{k-1}, y_j) and (x_k, y_{j+1}), in order to insure that the inverse mapping $(I + T)^{-1}$ exists (see Reference [32]). This second requirement translates as inequality constraints.

Since nodes on the boundary are allowed to move inside the domain, some computations will need the value of an image outside the domain D (e.g., $u \circ (I + T)^{-1}$ on D or the smoothing of u and v). To allow these computations, all the images are extended on $\mathbb{R}^2$ by assuming that there is no precipitation outside the domain, that is, $\forall (x, y) \notin D, \;\; u(x, y) = v(x, y) = 0$.

Appendix A.2. Minimization Method

In Reference [32], the minimization problem was solved by optimizing T_i one node after each other (for all nodes of the grid D_i). For each node, the coordinates x and y were updated alternatively using

a 1-D constrained non-linear optimization function. We use a different approach and optimize all the nodes at the same time. The solution does not depend on the order in which the nodes are optimized. However, the number of variables to optimize ($2 \cdot (2^i + 1)^2$ for grid D_i) increases exponentially with i. Constrained optimization algorithms become computationally expensive for such large numbers of independent variables. Thus, for computation efficiency, we used an iterative barrier approach ([34,35]).

In the barrier approach, the inequality constraints are added to the cost function J as penalization terms:

$$J_p(T_i) = J(T_i) + \beta \sum_h C_h(T_i) \tag{A1}$$

where C_h are constraint functions and β is the barrier coefficient. The constraints that cannot be written as bounds are converted into constraint functions C_h, such that $C_h(T_i) > 0$ if the constraint is violated and $C_h(T_i) = 0$ if it is respected.

This new minimization problem does not have inequality constraints, only bounds from our first requirement. Here, we minimized the penalized cost function J_p using the limited-memory quasi-Newton method for bound-constrained problems (L-BFGS-B) method (with the Python function scipy.optimize.minimize).

As mentioned above, the barrier approach is iterative. First, the cost function J_p is minimized using the L-BFGS-B method with $\beta = 1$. If all the constraints are respected, the procedure stops. Otherwise, the iterations continue with $\beta = 10 \cdot \beta$. The iterations continue until all the constraints are respected or until one of the stopping criteria is reached. We set two stopping criteria: (1) the decrease of the cost function P is smaller than $< 10^{-5}$ and (2) the root mean square difference of the grids D_i before and after is smaller $< 10^{-5}$.

Appendix B. Sensitivity Study for the Regulation Coefficients C_1, C_2 and C_3

The automatic registration is based on the minimization of the cost function J, which is composed of an observation term and a background term. The background term J_b consists of the three criteria defining the 'optimal' mapping T. The three criteria concern the magnitude (Equation (2)), the smoothness (Equation (3)) and the divergent (Equation (4)) of T. These three criteria can be tuned with the regulation coefficients C_1, C_2 and C_3 respectively. In this section, we look at the influence of these three coefficient on the automatic registration.

For this sensitivity study, we use the synthetic case described in Section 3.1. In the result section (Section 4), we empirically set the coefficients to $C_1 = 0.1$ ans $C_2 = C_3 = 1$. Here, we consider four more cases. We first look at the impact of the individual coefficients, setting the two other to zeros. Then we examine the influence of their intensity by multiplying them by five. The only difference between the results shown here and the ones in Section 3.1 are the regulation coefficients.

The impact of the coefficients on the morphed and warped signals can be seen in Table A1. It shows the mean absolute error (MAE) of both signals for the four cases described above and for the original case from Section 3.1. The MAEs are staying similar. The three cases with only one coefficient have a slightly higher MAE for the morphed signal than the original case, but slightly lower for the warped signal. Multiplying the three coefficients by five have the inverse effect. The MAE of the morphed signal is lower but that of the warped signal is slightly higher. The coefficients seem to have a limited impact on the warped and morphed signal.

Table A1. MAE of the warped (u_{warp}) and morphed u_{morp} signals compared to the target field (v) for the different sensitivity runs (for $I = 4$).

	Original	Only C_1	Only C_2	Only C_3	All Coef. $\times 5$
C_1	0.1	0.1	0	0	0.5
C_2	1	0	1	0	5
C_3	1	0	0	1	5
- MAE u_{morph}(mm/h)	0.0496	0.0520	0.0535	0.0515	0.0396
- MAE u_{warp}(mm/h)	0.1371	0.1363	0.1368	0.1366	0.1374

We now look at the impact of these coefficients on the mapping T. Figures A1–A4 show the mapping T and the resulting grid deformation for the four new cases. They can be compare to Figure 7 for the original case. The main differences occur in area without precipitation. Without the smoothing constraints, larger distortion occurs near the boundary near the lower-right corner (visible in Figures A1 and A3 but not in Figure A2). Similarly, the grid appears very stretched on the right upper side. The case with only the magnitude constraint ($C_1 = 0.1$) exhibit the smooth transformation (Figure A1). The divergent-free condition favours rotation-like patterns, in order to conserve the grid cell's volume. The weight of the coefficient also have an impact on the mapping T. By comparing Figures 7a and A4a, one can see that the distortion is similar over the rainy area, while the spatial displacement is smaller over the non-rainy area. The distortion is smoother and very similar to the regular one in the upper-left corner, i.e away from the rainy peak.

To summarize, without the three weak constraints the distorted grid shows unnatural distortion. However, the regulation coefficients mainly affect the mapping T in areas without rainfall. This is reflected in the small influence it has over the morphed or warped field (Table A1).

| (a) | (b) |
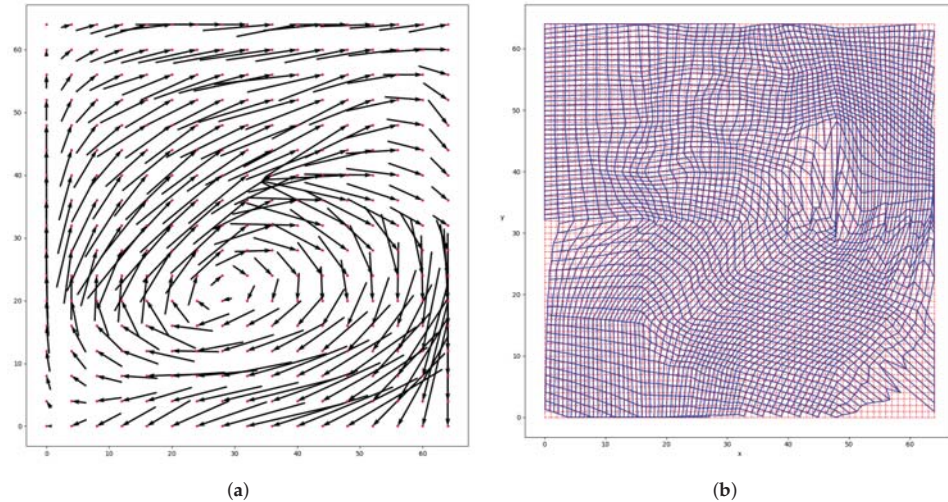
Figure A1. Sensitivity study: $C_1 = 0.1$ and $C_2 = C_3 = 0$. (**a**) Mapping T obtained from the automatic registration and (**b**) its effect on the pixel grid. (**a**) Mapping T. (**b**) Pixel grid D_n before (in red) and after (in blue) distortion by the mapping T.

(**a**)   (**b**)
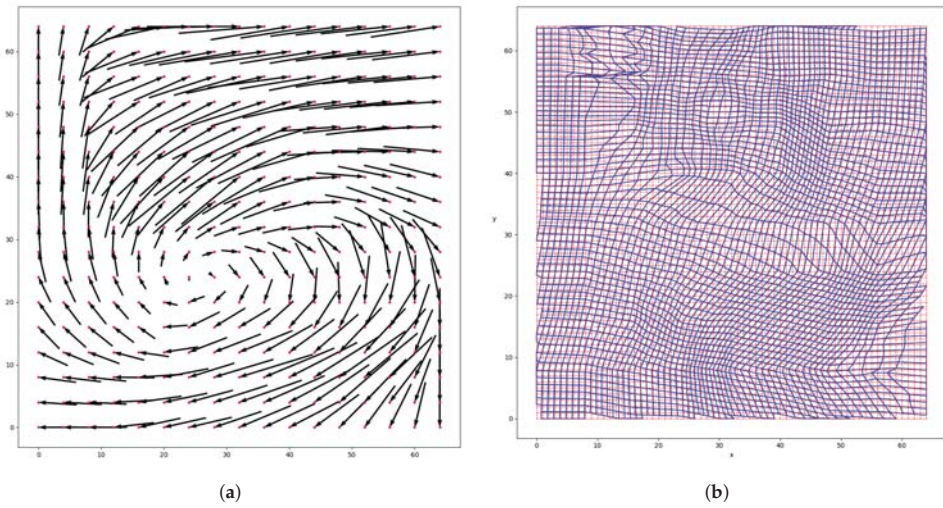
Figure A2. Sensitivity study: $C_2 = 1$ and $C_1 = C_3 = 0$. (**a**) Mapping T obtained from the automatic registration and (**b**) its effect on the pixel grid. (**a**) Mapping T. ((**b**) Pixel grid D_n before (in red) and after (in blue) distortion by the mapping T.).

(**a**)   (**b**)
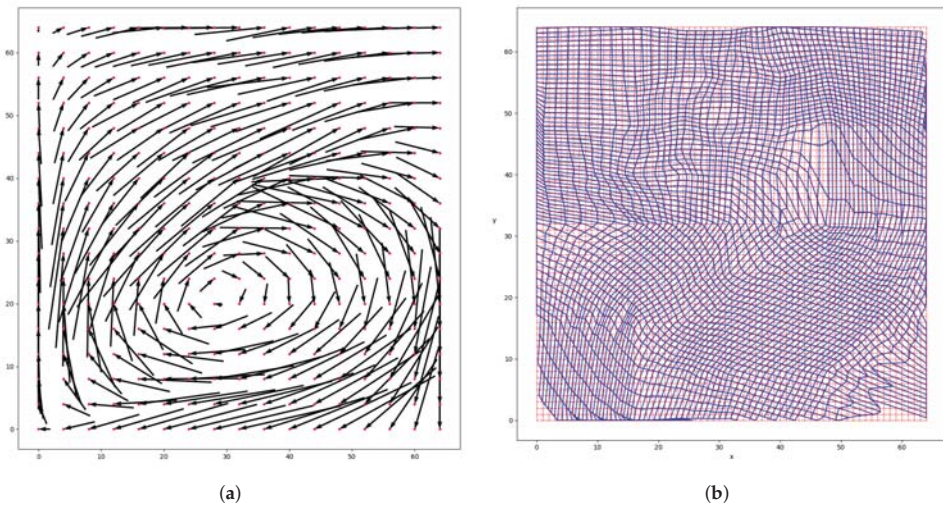
Figure A3. Sensitivity study: $C_3 = 1$ and $C_1 = C_2 = 0$. (**a**) Mapping T obtained from the automatic registration and (**b**) its effect on the pixel grid. (**a**) Mapping T. (**b**) Pixel grid D_n before (in red) and after (in blue) distortion by the mapping T.

(**a**) (**b**)
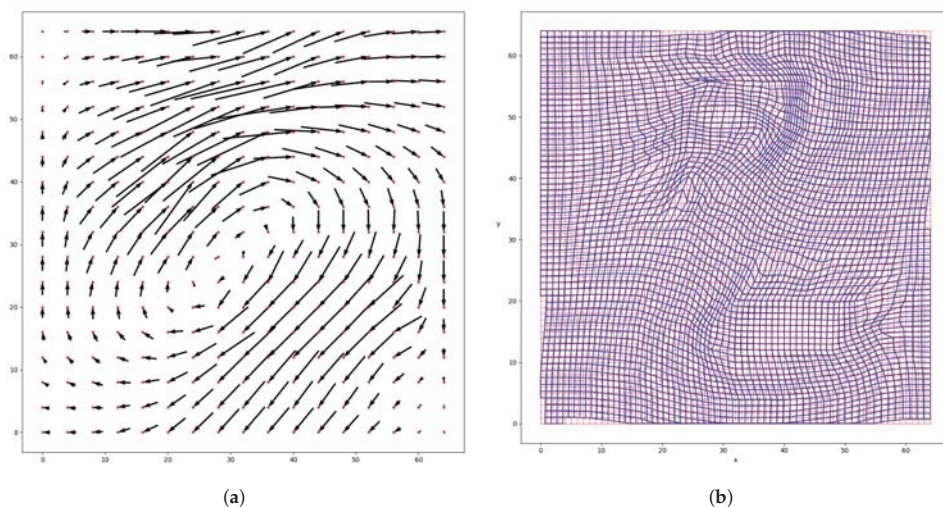
Figure A4. Sensitivity study: $C_1 = 0.5$ and $C_1 = C_2 = 5$. (**a**) Mapping T obtained from the automatic registration and (**b**) its effect on the pixel grid. (**a**) Mapping T. (**b**) Pixel grid D_n before (in red) and after (in blue) distortion by the mapping T.

References

1. Engel, T.; Fink, A.H.; Knippertz, P.; Pante, G.; Bliefernicht, J. Extreme Precipitation in the West African Cities of Dakar and Ouagadougou: Atmospheric Dynamics and Implications for Flood Risk Assessments. *J. Hydrometeorol.* **2017**, *18*, 2937–2957. [CrossRef]
2. Maranan, M.; Fink, A.H.; Knippertz, P. Rainfall types over southern West Africa: Objective identification, climatology and synoptic environment. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 1628–1648. [CrossRef]
3. Lorenz, C.; Kunstmann, H. The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis. *J. Hydrometeorol.* **2012**, *13*, 1397–1420. [CrossRef]
4. Morin, E.; Goodrich, D.C.; Maddox, R.A.; Gao, X.; Gupta, H.V.; Sorooshian, S. Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Adv. Water Resour.* **2006**, *29*, 843–860. [CrossRef]
5. Cristiano, E.; ten Veldhuis, M.C.; van de Giesen, N. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—A review. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3859–3878. [CrossRef]
6. Arnaud, P.; Bouvier, C.; Cisneros, L.; Dominguez, R. Influence of rainfall spatial variability on flood prediction. *J. Hydrol.* **2002**, *260*, 216–230. [CrossRef]
7. Xiao, Q.; Zou, X.; Kuo, Y.H. Incorporating the SSM/I-Derived Precipitable Water and Rainfall Rate into a Numerical Model: A Case Study for the ERICA IOP-4 Cyclone. *Mon. Weather. Rev.* **2000**, *128*, 87–108. [CrossRef]
8. Lopez, P. Direct 4D-Var Assimilation of NCEP Stage IV Radar and Gauge Precipitation Data at ECMWF. *Mon. Weather. Rev.* **2011**, *139*, 2098–2116. [CrossRef]
9. Gilleland, E.; Ahijevych, D.; Brown, B.G.; Casati, B.; Ebert, E.E. Intercomparison of Spatial Forecast Verification Methods. *Weather. Forecast.* **2009**, *24*, 1416–1430. [CrossRef]
10. Ahijevych, D.; Gilleland, E.; Brown, B.G.; Ebert, E.E. Application of Spatial Verification Methods to Idealized and NWP-Gridded Precipitation Forecasts. *Weather. Forecast.* **2009**, *24*, 1485–1497. [CrossRef]
11. Briggs, W.M.; Levine, R.A. Wavelets and Field Forecast Verification. *Mon. Weather. Rev.* **1997**, *125*, 1329–1341. [CrossRef]
12. Casati, B.; Ross, G.; Stephenson, D.B. A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl.* **2004**, *11*, 141–154. [CrossRef]

13. Casati, B. New Developments of the Intensity-Scale Technique within the Spatial Verification Methods Intercomparison Project. *Weather. Forecast.* **2010**, *25*, 113–143. [CrossRef]

14. Ebert, E.; McBride, J. Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrol.* **2000**, *239*, 179–202. [CrossRef]

15. Davis, C.; Brown, B.; Bullock, R. Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Mon. Weather. Rev.* **2006**, *134*, 1772–1784. [CrossRef]

16. Marzban, C.; Sandgathe, S. Cluster Analysis for Verification of Precipitation Fields. *Weather. Forecast.* **2006**, *21*, 824–838. [CrossRef]

17. Hoffman, R.N.; Liu, Z.; Louis, J.F.; Grassoti, C. Distortion Representation of Forecast Errors. *Mon. Weather. Rev.* **1995**, *123*, 2758–2770. [CrossRef]

18. Hoffman, R.N.; Grassotti, C. A Technique for Assimilating SSM/I Observations of Marine Atmospheric Storms: Tests with ECMWF Analyses. *J. Appl. Meteorol.* **1996**, *35*, 1177–1188. [CrossRef]

19. Grassotti, C.; Iskenderian, H.; Hoffman, R.N. Calibration and Alignment. *J. Appl. Meteorol.* **1999**, *38*, 677–695. [CrossRef]

20. Nehrkorn, T.; Woods, B.K.; Hoffman, R.N.; Auligné, T. Correcting for Position Errors in Variational Data Assimilation. *Mon. Weather. Rev.* **2015**, *143*, 1368–1381. [CrossRef]

21. Aonashi, K.; Eito, H. Displaced Ensemble Variational Assimilation Method to Incorporate Microwave Imager Brightness Temperatures into a Cloud-resolving Model. *J. Meteorol. Soc. Jpn. Ser. II* **2011**, *89*, 175–194. [CrossRef]

22. Nehrkorn, T.; Woods, B.; Auligné, T.; Hoffman, R.N. Application of Feature Calibration and Alignment to High-Resolution Analysis: Examples Using Observations Sensitive to Cloud and Water Vapor. *Mon. Weather. Rev.* **2014**, *142*, 686–702. [CrossRef]

23. Brewster, K.A. Phase-Correcting Data Assimilation and Application to Storm-Scale Numerical Weather Prediction. Part I: Method Description and Simulation Testing. *Mon. Weather Rev.* **2003**, *131*, 480–492. [CrossRef]

24. Lawson, W.G.; Hansen, J.A. Alignment Error Models and Ensemble-Based Data Assimilation. *Mon. Weather Rev.* **2005**, *133*, 1687–1709. [CrossRef]

25. Ravela, S.; Emanuel, K.; McLaughlin, D. Data assimilation by field alignment. *Phys. D Nonlinear Phenom.* **2007**, *230*, 127–145. [CrossRef]

26. Keil, C.; Craig, G.C. A Displacement-Based Error Measure Applied in a Regional Ensemble Forecasting System. *Mon. Weather Rev.* **2007**, *135*, 3248–3259. [CrossRef]

27. Keil, C.; Craig, G.C. A Displacement and Amplitude Score Employing an Optical Flow Technique. *Weather Forecast.* **2009**, *24*, 1297–1308. [CrossRef]

28. Marzban, C.; Sandgathe, S. Optical Flow for Verification. *Weather Forecast.* **2010**, *25*, 1479–1494. [CrossRef]

29. Marzban, C.; Sandgathe, S.; Lyons, H.; Lederer, N. Three Spatial Verification Techniques: Cluster Analysis, Variogram, and Optical Flow. *Weather Forecast.* **2009**, *24*, 1457–1471. [CrossRef]

30. Han, F.; Szunyogh, I. A Morphing-Based Technique for the Verification of Precipitation Forecasts. *Mon. Weather Rev.* **2016**, *144*, 295–313. [CrossRef]

31. Alexander, G.D.; Weinman, J.A.; Schols, J.L. The Use of Digital Warping of Microwave Integrated Water Vapor Imagery to Improve Forecasts of Marine Extratropical Cyclones. *Mon. Weather Rev.* **1998**, *126*, 1469–1496. [CrossRef]

32. Beezley, J.D.; Mandel, J. Morphing ensemble Kalman filters. *Tellus A* **2008**, *60*, 131–140. [CrossRef]

33. Mandel, J.; Beezley, J.D.; Eben, K.; Jurus, P.; Kondratenko, V.Y.; Resler, J. *Data Assimilation by Morphing Fast Fourier Transform Ensemble Kalman Filter for Precipitation Forecasts Using Radar Images*; UCD/CCM Report 289; Center for Computational Mathematics, University of Colorado at Denver: Denver, CO, USA, April 2010.

34. Jongen, H.T.; Meer, K.; Triesch, E. Penalty-, Barrier-, Multiplier-, Interior Point-Methods. In *Optimization Theory*; Springer: Boston, MA, USA, 2004; pp. 153–172. [CrossRef]

35. Nocedal, J.; Wright, S.J. Penalty and Augmented Lagrangian Methods. In *Numerical Optimization*; Springer: New York, NY, USA, 2006; pp. 497–528. [CrossRef]

36. Dezfuli, A.K.; Ichoku, C.M.; Huffman, G.J.; Mohr, K.I.; Selker, J.S.; van de Giesen, N.; Hochreutener, R.; Annor, F.O. Validation of IMERG Precipitation in Africa. *J. Hydrometeorol.* **2017**, *18*, 2817–2825. [CrossRef]

37. Dezfuli, A.K.; Ichoku, C.M.; Mohr, K.I.; Huffman, G.J. Precipitation Characteristics in West and East Africa from Satellite and in Situ Observations. *J. Hydrometeorol.* **2017**, *18*, 1799–1805. [CrossRef]

38. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Kidd, C.; Nelkin, E.J.; Sorooshian, S.; Tan, J.; Xie, P. *NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG)*; Algorithm Theoretical Basis Document Version 5.2; NASA GSFC: Greenbelt, MD, USA, 2018. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2_0.pdf (accessed on 16 September 2019).
39. Van de Giesen, N.; Hut, R.; Selker, J. The Trans-African Hydro-Meteorological Observatory (TAHMO). *Wiley Interdiscip. Rev. Water* **2014**, *1*, 341–348. [CrossRef]
40. Soenario, I.; Sluiter, R. *Optimization of Rainfall Interpolation*; Technical Report; Ministerie van Verkeer en Waterstaat, Konijklijk Nederlands Meteorologisch Instituut: De Bilt, The Netherlands, 2010.
41. Thiemig, V.; Rojas, R.; Zambrano-Bigiarini, M.; Roo, A.D. Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. *J. Hydrol.* **2013**, *499*, 324–338. [CrossRef]
42. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.J.M.; Weedon, G.P.; Brocca, L.; Pappenberger, F.; Huffman, G.J.; Wood, E.F. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6201–6217. [CrossRef]

Performance Assessment of SM2RAIN-CCI and SM2RAIN-ASCAT Precipitation Products over Pakistan

Khalil Ur Rahman [1], Songhao Shang [1,*], Muhammad Shahid [1,2] and Yeqiang Wen [1]

[1] State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, China
[2] NICE, SCEE, National University of Sciences & Technology Islamabad, Islamabad 44000, Pakistan
* Correspondence: shangsh@tsinghua.edu.cn; Tel.: +86-10-6279-6674

Abstract: Accurate estimation of precipitation from satellite precipitation products (PPs) over the complex topography and diverse climate of Pakistan with limited rain gauges (RGs) is an arduous task. In the current study, we assessed the performance of two PPs estimated from soil moisture (SM) using the SM2RAIN algorithm, SM2RAIN-CCI and SM2RAIN-ASCAT, on the daily scale across Pakistan during the periods 2000–2015 and 2007–2015, respectively. Several statistical metrics, i.e., Bias, unbiased root mean square error (ubRMSE), Theil's U, and the modified Kling–Gupta efficiency (KGE) score, and four categorical metrics, i.e., probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and Bias score, were used to evaluate these two PPs against 102 RGs observations across four distinct climate regions, i.e., glacial, humid, arid and hyper-arid regions. Total mean square error (MSE) is decomposed into systematic (MSEs) and random (MSEr) error components. Moreover, the Tropical Rainfall Measurement Mission Multi-Satellite Precipitation Analysis (TRMM TMPA 3B42v7) was used to assess the performance of SM2RAIN-based products at 0.25° scale during 2007–2015. Results shows that SM2RAIN-based product highly underestimated precipitation in north-east and hydraulically developed areas of the humid region. Maximum underestimation for SM2RAIN-CCI and SM2RIAN-ASCAT were 58.04% and 42.36%, respectively. Precipitation was also underestimated in mountainous areas of glacial and humid regions with maximum underestimations of 43.16% and 34.60% for SM2RAIN-CCI. Precipitation was overestimated along the coast of Arabian Sea in the hyper-arid region with maximum overestimations for SM2RAIN-CCI (SM2RAIN-ASCAT) of 59.59% (52.35%). Higher ubRMSE was observed in the vicinity of hydraulically developed areas. Theil's U depicted higher accuracy in the arid region with values of 0.23 (SM2RAIN-CCI) and 0.15 (SM2RAIN-ASCAT). Systematic error components have larger contribution than random error components. Overall, SM2RAIN-ASCAT dominates SM2RAIN-CCI across all climate regions, with average percentage improvements in bias (27.01% in humid, 5.94% in arid, and 6.05% in hyper-arid), ubRMSE (19.61% in humid, 20.16% in arid, and 25.56% in hyper-arid), Theil's U (9.80% in humid, 28.80% in arid, and 26.83% in hyper-arid), MSEs (24.55% in humid, 13.83% in arid, and 8.22% in hyper-arid), MSEr (19.41% in humid, 29.20% in arid, and 24.14% in hyper-arid) and KGE score (5.26% in humid, 28.12% in arid, and 24.72% in hyper-arid). Higher uncertainties were depicted in heavy and intense precipitation seasons, i.e., monsoon and pre-monsoon. Average values of statistical metrics during monsoon season for SM2RAIN-CCI (SM2RAIN-ASCAT) were 20.90% (17.82%), 10.52 mm/day (8.61 mm/day), 0.47 (0.43), and 0.47 (0.55) for bias, ubRMSE, Theil's U, and KGE score, respectively. TMPA outperformed SM2RAIN-based products across all climate regions. SM2RAIN-based datasets are recommended for agricultural water management, irrigation scheduling, flood simulation and early flood warning system (EFWS), drought monitoring, groundwater modeling, and rainwater harvesting, and vegetation and crop monitoring in plain areas of the arid region.

1. Introduction

Precipitation is one of the most critical components of global energy and water cycles and ranked first by the Global Climate Observing System (GCOS) [1,2]. Reliable long-term temporal precipitation records at fine spatiotemporal (<20 km at daily and sub-daily scales) resolution is crucial for planning and managing water resources, drought assessment, flood forecasting, assessment of crop water requirements, hydrometeorology, and climate studies [3–7]. Precipitation also plays an important role in weather prediction, agricultural management, vector and water-borne diseases [8].

Precipitation is highly erratic spatiotemporally, which makes its estimation challenging both with ground observation (rain gauges and meteorological radars) and satellite precipitation products (PPs). The complex topography, varying climate, dense vegetation, and coastal regions attribute to further complexity in precipitation estimation [9,10]. Ground-based rain gauges (RGs) provide accurate local precipitation estimates [11], and are considered as the most reliable precipitation record source for hydrological modeling and monitoring purposes. However, their non-homogeneous coverage, limited spatial representativeness, and high maintenance cost restrain their global scale application [12].

Meteorological models, such as reanalysis products, are alternatively used to estimate precipitation, especially in regions with scarce RGs and reliable ground observations [13]. The uncertainties associated with these estimates in areas with scarce RGs can be substantial [14]. Therefore, in the past 30 years, different remote sensing techniques have been developed to improve the precipitation estimations and provide full regional and global spatial coverage [15]. The standard precipitation measurement methods are based on instantaneous precipitation measurements from space retrieved from microwave radiometers, radars, and infrared sensors [16]. These methods invert the atmospheric signals reflected or radiated by hydrometeors and known as the "top-down" approach [9].

PPs estimate precipitation from thermal infrared sensors onboard geosynchronous earth orbit (GEO), and microwave sensors (both passive and active) onboard low-earth orbit (LEO) satellites [10,17]. Some PPs combine infrared- and microwave-based estimates by utilizing high temporal resolution infrared platforms and shows high precision in precipitation estimation of microwave sensors [18]. Such PPs include the most recently developed PPs with high spatial (0.1°) and temporal (30 minutes) resolution, i.e., Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) of the Global Precipitation Measurement (GPM) mission [15], near-real-time and post real versions of the Tropical Rainfall Measurement Mission Multi-Satellite Precipitation Analysis (TRMM TMPA 3B42RT, 3B42V6/V7) [19], the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) [20], Climate Prediction Center (CPC) Morphing (CMORPH) [21], Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [22,23], Global Satellite Mapping of Precipitation (GSMaP) [24] and many more.

Soil Moisture (SM) TO RAIN (SM2RAIN) is an algorithm based on approaches provided by Crow et al. [25] and Pellarin et al. [26] and further developed by Brocca et al. [27], which can be used for direct precipitation estimation from in-situ and/or satellite-based SM observations. The method provides daily precipitation estimates on a global scale and has been successfully applied to satellite SM data [9,28,29]. SM2RAIN approach is appropriate for accumulated precipitation estimates and has been verified against in-situ SM data and single-sensor SM products [30,31] providing reliable results at the regional scale [32–34].

SM2RAIN is based on the "bottom-up" approach that estimates precipitation from various single-sensors surface soil moisture (SSM), as opposed to "top-down" approach of other PPs [27,30, 35,36]. The bottom-up approach differs from top-down approach in such a way that the bottom-up approach considers accumulated precipitation while the top-down approach is based on instantaneous

precipitation rate. The bottom-up approach has an edge over top-down when accumulated precipitation (e.g., daily precipitation) estimates are desirable as this approach requires a limited number of satellite sensors and measurements. On the other hand, the limitations of the bottom-up approach are its dependence on topography (low accuracy over complex topography) and land use (low accuracy over dense vegetation), unable to estimate precipitation in over-saturated soil, and applicable only to terrestrial rainfall [9].

The bottom-up approach has been extensively validated over extended-spectrum of spatial scales, including global [9,25,37], continental [35,38], and regional [32,39–42] scales. Different SM PPs are considered, such as SMOS (Soil Moisture Ocean Salinity Mission, [35]), ASCAT (Advanced SCATterometer, [43]), AMSR-E (Advanced Microwave Scanning Radiometer, [44]), and SMAP (Soil Moisture Active and Passive, [28,29]). Recently, a number of studies have employed precipitation estimates from satellites obtained through the bottom-up approach in hydrological and water resources applications [3,17,31,45]. These studies demonstrated the potential benefits and main limitations of the bottom-up approach in estimating precipitation from space. The accuracy of the bottom-up approach is strongly influenced by the accuracy and temporal resolution of satellite SM products [46].

The objectives of this study were two-fold, i.e., to assess the performance of SM2RAIN-CCI and SM2RAIN-ASCAT for the first time in Pakistan, a country with a complex topography and diverse climate, and to evaluate the performance of SM2RAIN-derived products against extensively evaluated PP in Pakistan (TRMM TMPA 3B42v7, hereinafter referred as TMPA). The analysis was performed in four climate regions of Pakistan, i.e., glacial, humid, arid and hyper-arid regions, during the period 2000-2015. This study is the first of its kind that evaluates the quality of SM2RAIN-based precipitation datasets across Pakistan. This study can be used as a reference for hydrological modeling, water resources management, and agricultural water management practices. Moreover, there are very limited studies on integrated performance evaluation of SM2RAIN-based precipitation datasets in regions with complex topography and diverse climate, especially in Pakistan.

Pakistan is a developing country that has very limited/scarce rain gauges (RGs), which are non-homogenously distributed. Even with a significant increase in the number of RGs over the last few decades, their density does not meet scientific and practical requirements. To overcome the current scenario, Pakistan needs the application of advanced remote sensing techniques for hydrological and meteorological applications, climate change studies, agricultural water management, and water resources management. This study evaluated the soil moisture-based precipitation datasets, which will be a useful alternative to conventional precipitation products for different hydrological and meteorological applications and will address the data scarcity problems in Pakistan.

The significance of the current study is: (1) This study demonstrates the worth/quality of SM2RAIN-based datasets by evaluation on spatial and temporal scales, (2) the spatial and temporal evaluation helps us to understand where, when and how these precipitation products might be used, (3) better performance of these products in the arid (and semi-arid regions), i.e., Punjab province of Pakistan considered as agricultural hub of the country, illustrates the potential application of SM2RAIN-based datasets for agriculture (vegetation and crop growth monitoring) and agricultural water management, and much more.

2. Materials and Methods

2.1. Study Area

Pakistan lies between latitude 23.5°–37.5°N and longitude 62°–75°E in the western region of South-Asia with an area of 803,940 km^2 (Figure 1a). Pakistan is bounded by China at its north, Iran and Afghanistan at the west, India at the east, and the Arabian Sea at the south. The study area has complex topography with a diverse climate. Elevation of Pakistan ranges from 8600 m (at Hindukush-Himalaya mountain ranges) at the extreme north to 0 m (at the coast of Arabian Sea) [47]. The climate of Pakistan

changes abruptly, ranging from glacial to humid, arid and hyper-arid regions. According to climate variability, the study area is divided into glacial, humid, arid and hyper-arid climate regions (Figure 1b).
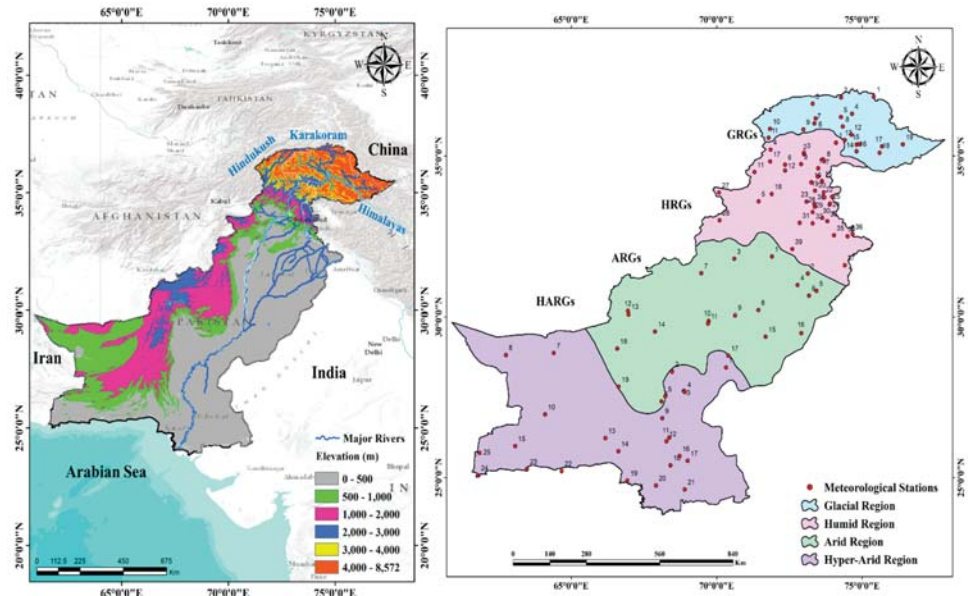
Figure 1. Study area and geographical location of rain gauges (RGs). (**a**) The elevation map of Pakistan derived from Shuttle Radar Topography Model (SRTM), (**b**) The four climate regions considered in this study.

The glacial region, situated in the extreme north of Pakistan, is characterized by permanent glaciers and snow cover. Mean elevation of the glacial region is 4158 m, with the mean annual precipitation of 348 mm. Snow and glaciers from the glacial region melt in summer and feed the Indus river and its branches. Water from snowmelt is the primary source of agricultural, domestic, and industrial water use in Pakistan. Moreover, the country is threatened by excessive snow and glacier melt that causes acute flooding, such as the 2010 flood which severely damaged the infrastructure, the country's economy and took thousands of lives. The glacial region is located between 34°N to 38°N and comprised of the world-famous mountains of Hindukush Himalayas, which are famous for snow after the polar regions.

The humid region consists of mountains ranges of Hindukush–Karakoram–Himalaya (HKH) ranges, and all of the main rivers originating from these mountains, including Indus, Kabul, Gilgit, Hunza, Swat, Panjkora, Kurram, and Jhelum rivers. Mean elevation and mean annual precipitation of the humid region is 1286 m and 852 mm, respectively. The humid region is the hydraulically developed region of the country which consists of the largest dam, Tarebla dam, and Mangla dam. The Tarbela dam, constructed on the Indus river, has 3500 MW of capacity, while the Mangla dam constructed on Jhelum river has 1000 MW hydropower capacity [48]. Moreover, the humid region is built up with extensive barrages, headworks, and a developed integrated irrigation canal system. The secondary purpose of the hydraulic structures is to support the agriculture sector of the country by meeting the extensive irrigation water requirements [48].

The arid region is mostly comprised of the major agricultural regions of Punjab province. The Indus river and other rivers flowing through this region are the primary sources of water for agricultural, domestic, and industrial purposes. Mean elevation of the arid region is 663 m with mean annual precipitation of 322 mm. Extreme west (elevated mountainous areas) of this region is cold in winter

(with snowfall in December and January) and hot in the summer seasons. The remaining of the arid zone is dry and hot in nature where maximum temperatures are recorded in Sibbi and Jacobabad. The area is plain with some area included in the famous Thar Desert [49].

Sindh and Balochistan provinces and the south part of the Punjab province of Pakistan lie in the hyper-arid region. Most of the hyper-arid region is comprised of deserts, plateaus, barren lands, dry mountains, and coastal region along the Arabian Sea. Mean elevation and mean annual precipitation in the hyper-arid region are 444 m and 133 mm, respectively.

Precipitation in Pakistan is spatially varying from the maximum of 1500 mm/a in the north to a minimum of 100 mm/a in the south. Monsoon and winter (western disturbance) are the dominant seasons where maximum rainfall occurs [50]. Monsoon (July to September) precipitation in Pakistan commences from the Bay of Bengal entering Pakistan from northeast and east sides. Heavy and intense precipitation accounting 55%–60% of precipitation per annum occurs in Pakistan during the monsoon season [51]. Winter (December to March) precipitation originating from the Mediterranean Sea enters Pakistan through Afghanistan and Iran. Only 30% of total precipitation (moderate magnitude) occurs during winter. Precipitation across Pakistan varies spatially in magnitude; low precipitation in the glacial region from 34 to 36° N (<100 mm/month), high in the humid region from 29 to 33° N (>700 mm/month in North-East) and low again in the hyper-arid region from 24 to 28° N (around 100 mm) [47].

2.2. Ground-Based Precipitation Data

Rain gauges (RGs) are considered as standard sources of precipitation estimates that are used for satellite precipitation calibration and validation processes. The Pakistan Meteorology Department (PMD) and the Snow and Ice Hydrology Project (SIHP) of Water And Power Development Authority (WAPDA) own the climate and hydrological data in Pakistan. SIHP operates the meteorological stations at high elevation mostly located in glacial and humid regions. For the current study, daily precipitation records from 102 meteorological stations for the period 2000–2015 were collected from both organizations, where precipitation records of 79 RGs were collected from PMD and 23 from WAPDA. The location of meteorological stations across Pakistan is shown in Figure 1b. RGs in the current studies are named with respect to each climate region, i.e., RGs in glacial, humid, arid and hyper-arid regions are represented as GRG, HRG, ARG, and HARG, respectively. Main features and number of RGs in selected climate regions are presented in Table 1.

Table 1. Main features of the selected four climate regions.

Climate Region	Area (km2)	Mean Elevation (m)	Mean Annual Precipitation (mm)	Meteorological Stations
Glacial	72,774	4158	348	19
Humid	137,753	1286	852	39
Arid	270,484	633	322	19
Hyper-Arid	322,929	444	133	25

The data obtained from PMD and WAPDA were manually collected, and consequently might be subjected to personal and instrumental errors. Besides, the stations located at high elevation may also be subjected to splashing and wind errors. These associated errors affect the quality of the data. Therefore, PMD and WAPDA correct the RGs precipitation following the World Meteorology Organization's standard code (WMO-N). Moreover, the quality of the data was tested using skewness and kurtosis methods, and missing data was filled using the zero-order method [47].

2.3. Satellite-Based Precipitation Products

2.3.1. SM2RAIN-CCI

The European Space Agency Climate Change Initiative (ESA-CCI) released the ESA CCI SM (Soil Moisture) v3.2 datasets in early 2017. SM retrievals from active and passive microwave instruments onboard different satellite platforms were merged to develop the dataset, which provides global daily SM estimates with 0.25° spatial resolution from 1978–2015 [52–54]. Accumulated precipitation estimates between 00:00 to 23:59 UTC are obtained from an application of a weighted average-based integration procedure. Quality flag provided within the raw SM observations (i.e., EA CCI SM v3.2) is used to remove the low-quality data and observations distinguished by retrieval issues (e.g., glacial regions, frozen soil, complex topography, and dense vegetation). The SM2RAIN algorithm was calibrated on a pixel-by-pixel basis during the periods of 1998–2001, 2002–2006, and 2007–2013 against Global Precipitation Climatology Centre Full Daily Data (GPCC-FDD) [37]. These different calibration periods are dependent on types of data and sensors that have been utilized in developing the active and passive SM datasets.

The SM2RAIN algorithm, a novel approach, estimates precipitation from SM using the inverted soil water balance equation [9]. The basic assumption of the SM2RAIN algorithm is that surface runoff and evapotranspiration rate are insignificant during the precipitation event [9,32]. The simple form of soil–water balance equation can be illustrated as

$$p(t) = Z^* \frac{ds(t)}{dt} + a \cdot s(t)^b \tag{1}$$

where $p(t)$ is precipitation rate [L/T], $Z^*=nZ$ is soil water capacity [L] in the soil layer with a depth of Z and porosity of n, $s(t)$ represents relative soil saturation, t is time [T], and the two parameters a and b represent the nonlinearity between drainage and soil saturation. The parameters Z^*, a, and b are calculated through calibration analysis [37]. The main limitation of the SM2RAIN algorithm is that it is not able to estimate the precipitation when the soil is saturated or nearly saturated, because the SM algorithm is unable to derive SM variation as the SM approaches saturation [30]. A more detailed description of SM2RAIN algorithm can be found in [32]. In the current study, v2.0 of the SM2RAIN-CCI product, available from 1 January 1998 to 31 December 2015 (released in July 2018) with 0.25° spatial and daily temporal resolutions, is used.

2.3.2. SM2RAIN-ASCAT

The SM2RAIN algorithm was applied to three surface SM products obtained from Advanced Microwave Scanning Radiometer for Earth Observing (AMSR-E), Soil Moisture and Ocean Salinity (SMOS), and Advanced SCATterometer (ASCAT) to develop the SM2RAIN-ASCAT dataset [9]. ASCAT is a scatterometer and consists of the space segment of the EUMETSAT Polar System (EPS), operating at 5.255 GHz (C-band VV polarization) onboard MetOp A, B, and C satellites [55]. The spatial resolution of SM2RAIN-ASCAT is the same as the parent ASCAT SM dataset (previously developed datasets were resampled to spatial resolutions of 0.5° and 1°). SM2RAIN-ASCAT has 12.5 km spatial and daily temporal resolutions available for the period 2007–2018 [46]. The SM2RAIN-ASCAT showed more accuracy when evaluated on the basis of root mean square error (RMSE), Pearson correlation coefficient (R), and detection of precipitation [9].

The main differences between SM2RAIN-CCI and SM2RAIN-ASCAT datasets are the input SM datasets (input products for SM2RAIN-CCI are the ESA CCI SM product [56], and the time span of SM2RAIN-CCI that is available, namely, the period of 1998–2015). Moreover, the SM2RAIN-ASCAT is bias-corrected and calibrated against the Medium-Range Weather Forecasts (ERA5) instead of GPCC-FDD as SM2RAIN-CCI [43,46]. In the current study, version 1.0 of SM2RAIN-ASCAT, released in March 2019, is used.

2.3.3. TRMM TMPA

TRMM TMPA 3B42 v7 (hereinafter TMPA) was established by the joint efforts of the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and the Japan Aerospace Exploration Agency (JAXA). TMPA has 0.25° (~25 km at the equator) spatial resolution ranging from 50° S and 50° N bands with 3 hours temporal resolution available from January 1998 to the present. Precipitation estimation is based on the observations in infrared and microwave bands obtained by satellites. Re-analyzed precipitation data from the GPCC-FDD dataset was used to compute these multi-satellite precipitation estimates [57,58]. The complete description of TMPA is provided by reference [19]. The motivation behind the selection of TMPA is the fact that it has been evaluated comprehensively by many studies and is reported for its supremacy in comparison to other datasets [47,59].

2.4. Methods

Figure 2 represents the flow chart of the proposed methodology applied in the current study. SM2RAIN-CCI is compared and evaluated against the RGs during 2000–2015; while, SM2RAIN-ASCAT is evaluated during 2007–2015. Moreover, due to the difference in time duration of the considered PPs, the performance of each dataset is compared with TMPA during the shared period from 2007 to 2015. To ascertain the impartial inter-comparison of SM2RAIN-ASCAT with SM2RAIN-CCI and TMPA, SM2RAIN-ASCAT has been re-gridded using nearest neighbor algorithm at 0.25° spatial resolution. Figure 2 shows that inter-comparison of daily precipitation estimates from SM2RAIN-CCI, SM2RAIN-ASCAT, and TMPA with RGs is conducted to assess the spatial consistency of precipitation datasets over Pakistan. Further, the performance of precipitation datasets is also assessed based on precipitation topography and seasonal time scale using continuous and categorical metrics. Finally, systematic and random error components of SM2RAIN-CCI and SM2RAIN-ASCAT are also assessed utilizing a decomposition method (the final step in Figure 2).
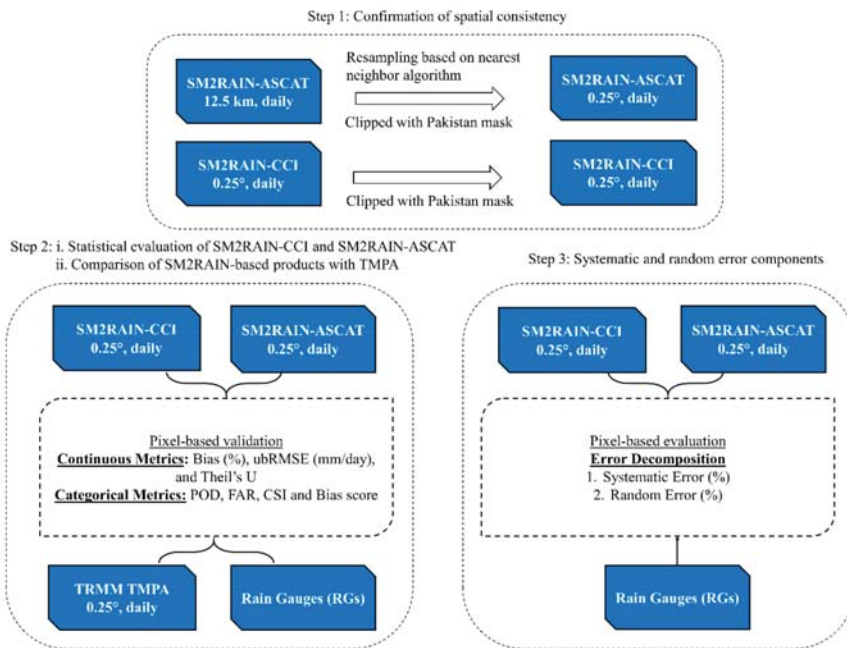

Figure 2. A general framework for the proposed methodology.

Three continuous metrics were used to assess the performance of SM2RAIN-CCI and SM2RAIN-ASCAT against the RGs. These metrics were based on inter-comparison of precipitation estimates from each product with precipitation records from RGs on a pixel scale. The Bias (B), unbiased root mean square error (ubRMSE), and Theil's U coefficient (U) are considered in this study (Table 2). B is used to estimate the over/under-estimation of precipitation by each product, where positive values indicate overestimation and negative values underestimation. ubRMSE measures the difference between estimated and observed precipitation values. ubRMSE attain only positive values where minimum values indicate better performance. Theil's U predicts the forecasting accuracy of PPs with respect to RGs. The lower bound of Theil's U is zero, indicating the perfect forecast; the value of 1 indicates that PPs forecast the same error as the naïve no-change extrapolation. A value greater than 1 indicates the worst forecasting and has to be rejected [60].

Table 2. Statistical metrics used to assess the performance of SM2RAIN-CCI and SM2RAIN-ASCAT. E is the estimated precipitation (PPs), O is the observed precipitation records from RGs, n is the total number of data points, $\overline{E}$ is the average of estimated precipitation, and $\overline{O}$ is the average of observed precipitation.

Statistical Metrics	Formula	Perfect Value
Bias	$B = \dfrac{\sum_{i=1}^{n}(E_i - O_i)}{\sum_{i=1}^{n} O_i}$	0
Root Mean Square Error	$\text{RMSE} = \sqrt{\dfrac{1}{n} \sum_{i=1}^{n}(E_i - O_i)^2}$	0
Unbiased root mean square error	$\text{ubRMSE} = \sqrt{\text{RMSE}^2 - B^2}$	0
Theil's U	$U = \sqrt{\dfrac{\sum_{i=1}^{n}(E_i - O_i)^2}{\sum_{i=1}^{n} E_i^2}}.$	0

Moreover, the method developed by AghaKouchak, et al. [61] is used to decompose the total mean square error (MSE) in PPs into random (MSEr) and systematic (MSEs) errors, which are presented as

$$\text{MSE} = \text{MSE}_r + \text{MSE}_s \tag{2}$$

$$\text{MSE}_r = \frac{1}{n} \sum_{i=1}^{n}\left(E_i - E_i^*\right)^2 \tag{3}$$

$$\text{MSE}_s = \frac{1}{n} \sum_{i=1}^{n}\left(E_i^* - O_i\right)^2 \tag{4}$$

$$E_i^* = a \times O_i + b \tag{5}$$

E_i^* (mm/day) in Equations (3)–(5) is calculated on the daily scale by least squared linearly regressing the PPs estimates against the RGs at each pixel. Therefore, a and b in Equation (5) are offset and scale parameters [62].

In addition, the performance of PPs was also assessed using the modified Kling–Gupta efficiency (KGE) score. KGE combines correlation coefficient (r), bias ratio (β) and variability ratio (γ) [63], which is represented as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{6}$$

$$r = \frac{\sum_{i=1}^{n}\left(E_i - \overline{E}\right)\left(O_i - \overline{O}\right)}{\sqrt{\sum_{i=1}^{n}\left(E_i - \overline{E}\right)^2}\sqrt{\sum_{i=1}^{n}\left(O - \overline{O}\right)^2}} \tag{7}$$

$$\beta = \frac{\overline{E}}{\overline{O}} \tag{8}$$

$$\gamma = \frac{(\text{CV})_E}{(\text{CV})_O}, \ (\text{CV})_E = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(E_i - \overline{E}\right)^2}/\overline{E}, \ (\text{CV})_O = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2}/\overline{O} \qquad (9)$$

where CV denotes the coefficient of variation. The perfect values of KGE, β and γ are all 1.

The precipitation detection competences of SM2RAIN-CCI and SM2RAIN-ASCAT are examined using four categorical metrics (Table 3), including the probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and Bias score (BS). POD represents the capability of PPs to detect precipitation events. A threshold of 1 mm is considered to distinguish precipitation from dry days (no precipitation) at any time scale [64]. FAR indicates the fraction of predicted precipitation event that did not occur. POD and FAR ranges between 0–1 with a perfect values of 1 and 0, respectively [62]. CSI is a fraction of precipitation events correctly detected by PPs. The CSI value ranges between 0 to 1, with the perfect value of 1 [65]. BS represents the fraction of all PPs precipitation events that were correctly predicted. The range of BS values is 0 to 1 with perfect score equal to 1.

Table 3. Categorical metrics, where *hit* indicates the number of precipitation events correctly detected both by PPs and RGs, *miss* indicates the number of precipitation events not detected by PPs but recorded by RGs, *false_alarm* is number of precipitation events detected by PPs while no precipitation records are available at RGs, and *N* is the total number of events.

Statistical Metric	Formula	Perfect Score
Probability of Detection	$\text{POD} = \frac{hit}{hit+miss}$	1
False_alarm_ratio	$\text{FAR} = \frac{false\_alarm}{hit+false\_alarm}$	0
Critical-Success-Index	$\text{CSI} = \frac{hit}{N}$	1
Bias score	$\text{BS} = \frac{hit+false\_alarm}{hit+miss}$	1

3. Results

3.1. Regional Evaluation of SM2RAIN-CCI Dataset

Accuracy of daily precipitation estimates from the SM2RAIN-CCI and SM2RAIN-ASCAT were assessed using the RGs data as a benchmark for each climate region during the period 2000–2015 and 2007–2015, respectively. The continuous metrics (listed in Table 2) values are statistical quantification of variations in the amount of precipitation from SM2RAIN-CCI/ASCAT datasets from the RGs data at the pixel scale (considered only the common pixels). The ordinary kriging method is used to interpolate the pixel base data over entire Pakistan (climate regions) to comprehensively understand the spatial distribution trend of error characteristics.

Figure 3 shows the spatial distribution of statistical metrics of SM2RAIN-CCI over four distinct climate regions, i.e., glacial, humid, arid and hyper-arid regions in Pakistan. Precipitation data for SM2RAIN-CCI is not available in the glacial region and high elevated areas of humid region due to masking, i.e., precipitation data of SM2RAIN-CCI is masked out in frozen soil, snow-dominated regions and complex topography.

Figure 3a represents the spatial distribution of SM2RAIN-CCI biases (B) across Pakistan in comparison with RG observations for the study period (2000–2015) on a daily temporal scale. In the humid region, larger negative biases (underestimation) are evident over north-east of the humid region but decreases gradually towards the west (slightly overestimated). The north-east of humid region consists of the Mangla dam, barrages, headworks and an extensively developed integrated irrigation canal system (comprised of 12 interlink canals and 43 independent irrigation canals commands) of the country (hereinafter called hydraulically developed areas), which results in the groundwater level rise in the vicinity of these structures and the soil water saturation in irrigated land after irrigation. Besides, the area is also subjected to heavy and intense precipitation that may saturate the soil in a quick span of time. Since the SM2RAIN-based products work on the principles of soil–water balance, the excessive precipitation cannot be considered when the soil gets saturated, and hence resulting in

higher biases (higher underestimation). The sign of biases shows an interesting feature along the west of arid region indicating a considerable underestimation of precipitation. Western side of arid regions is comprised of mild elevated mountainous ranges (i.e., Koh-e-Suleman, Koh-e-Chiltan, Koh-e-Murdaar, and Koh-e-Takatu) having cold weather in winter and hot in summer with mean annual precipitation of 317 mm [47]. Precipitation is underestimated in this region due to the low infiltration capacity (the area is mostly covered with rocks) and also due to the snow factor in the winter season. However, precipitation is overestimated in the plain agricultural areas of the arid region (east and south-east). Hyper-arid region is the hottest region located at the coast of the Arabian Sea and experience mean annual precipitation of 133 mm. Precipitation is overestimated in the region by SM2RAIN-CCI, which may be due to seawater intrusion from the Arabian Sea [66]. High overestimation is observed in the south-east coastal area which gradually decreases towards the west.

Figure 3b presents the spatial distribution of ubRMSE over different climate regions of Pakistan. A common trend of lower ubRMSE towards the west in comparison to the east of each climate region is observed in the figure. Higher ubRMSE is observed in the hydraulically developed areas in the north-east of humid region. Lower ubRMSE is observed in the hyper-arid region, most specifically near the coast of the Arabian Sea. The hyper-arid region is a precipitation deficit region where there is plenty of time for precipitation to infiltrate and saturate the soil. The reason for lower ubRMSE in the hyper-arid region is low intensity and low magnitude precipitation. Higher ubRMSE was observed at HRG23 (12.44 mm/day) while minimum at HARG8 (1.83 mm/day).

Figure 3c depicts the distribution of Theil's U coefficient that shows the accuracy of SM2RAIN-CCI to accurately detect a precipitation event. Smaller values (close to zero) represent better forecasting accuracy, while values closer to 1 depicts poor forecasting. The figure reveals better forecasting accuracies at ARG16 (0.24) and ARG8 (0.23), with poor forecasting accuracies in north-east of humid region, more specifically at HRG23 (0.66) and HRG24 (0.69), respectively. Locations of high accuracies (minimum Theil's U) in the arid region are in the vicinity of the agricultural region (ARG9) and the Thar desert (ARG16) where soil is mostly unsaturated. The only source of water for agriculture is irrigation water. Agriculture in the location helps to preserve the water in the soil. Beside the arid region, moderate accuracy is depicted in south-east of the hyper-arid region near HARG18, HARG19, HARG20, and HARG21. The average values of Theil's U in the humid, arid and hyper-arid regions were 0.51, 0.40, and 0.42, and the median values were 0.53, 0.43, and 0.41, respectively.

Spatial distributions of mean square error components, i.e., random and systematic errors, in SM2RAIN-CCI across Pakistan are presented in Figure 3d,e. Before the integration of any PP in the hydrological application, the knowledge about systematic and random errors for implementation of any statistical adjustment and bias correction is extremely vital [61]. In comparison with the random error component, systematic errors have larger contributions over Pakistan. The analysis depicts that SM2RAIN-CCI needs proper adjustments in biases before assimilating into any hydrological application. Larger systematic errors are observed in the hyper-arid and downstream portion of arid regions while minimum values are in the humid region.

Figure 3f presents the spatial distribution of KGE score for the SM2RAIN-CCI product as compared to RGs observation. Results shows smaller KGE scores in the humid region, which increases to maximum values in the middle of the arid region, then decreasing at the end of the arid region. Smaller KGE scores in north-east of the humid region might be due to the relatively poor performance of SM2RAIN-CCI in the hydraulically developed areas. Besides that, the uncertainties in gauge-based estimates associated with sparse gauge density also play a vital role. The highest KGE scores were observed at ARG3 (0.78), HRG11 (0.74), and ARG15 (0.73). The minimum KGE score was observed at HRG24 (0.13).

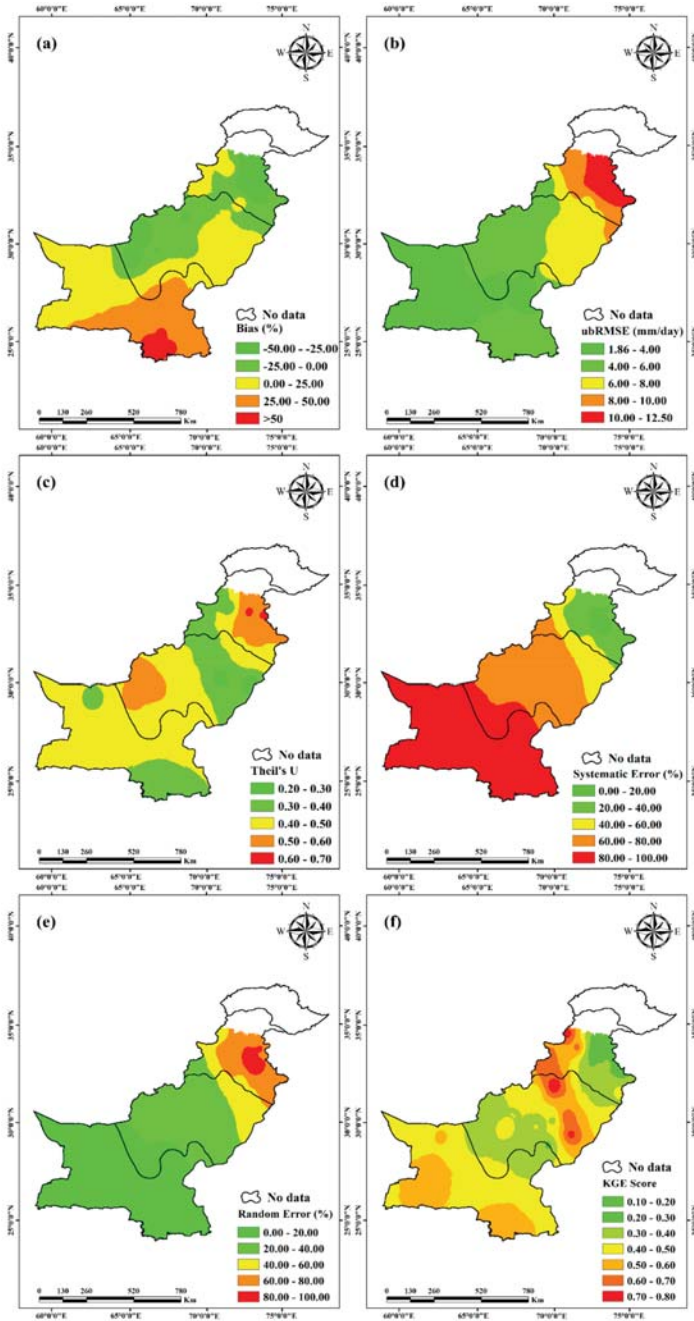**Figure 3.** Spatial distribution of bias (**a**), ubRMSE (**b**), Theil's U coefficient (**c**), systematic error (**d**), random error (**e**), and KGE score (**f**) based on a daily scale across Pakistan from SM2RAIN-CCI compared to RGs data for the period of 2000–2015.

3.2. Regional Evaluation of SM2RAIN-ASCAT Dataset

Figure 4 shows the spatial distribution of statistical metrics of SM2RAIN-ASCAT. The bias of SM2RAIN-ASCAT (Figure 4a) shows an almost similar trend compared to SM2RAIN-CCI (Figure 3a). However, the magnitude of biases at each RG is significantly reduced when compared to SM2RAIN-CCI. Precipitation is highly underestimated at the north-east sides of the glacial and humid regions of the country, which gradually decreases towards the west of the regions. The reason for high underestimation might be the abundant water availability in the specific region due to dams, barrages, headworks, and well-developed irrigation systems. Besides, snowmelt and high precipitation events that saturate the soil also contribute to an underestimation of precipitation in high elevated regions (i.e., the glacial region and upstream areas of the humid regions). SM2RAIN-ASCAT highly overestimated precipitation in the south-east of the hyper-arid region. The average (median) values of bias in the glacial, humid, arid and hyper-arid regions were −28.76% (−29.25%), −29.86% (−23.31%), −6.18% (−12.68%) and 34.87% (27.66%), respectively.

SM2RAIN-ASCAT depicts an improved performance in comparison with SM2RAIN-CCI considering ubRMSE (Figure 4b). ubRMSE shows the same spatial trend as compared to SM2RAIN-CCI with maximum and minimum values in the humid and hyper-arid regions, respectively. The regional average (median) ubRMSE values were 3.73 mm/day (3.57 mm/day) in glacial, 7.87 mm/day (7.54 mm/day) in humid, 4.83 mm/day (4.32 mm/day) in humid and 2.21 mm/day (2.19 mm/day) in hyper-arid regions. Maximum and minimum ubRMSE values were depicted by HRG12 (9.91 mm/day) and HARG8 (1.21 mm/day).

SM2RAIN-ASCAT shows high forecasting accuracy at the south-east of the arid region while moderate accuracies in extreme south of hyper-arid region (Figure 4c). Lower accuracies are depicted in humid region mostly in the vicinity of hydraulically developed areas. The regional average (median) values of Theil's U coefficient were 0.40 (0.39) in glacial, 0.46 (0.48) in humid, 0.25 (0.19) in arid, and 0.30 (0.28) in hyper-arid regions. The maximum and minimum of Theil's U values were 0.59 (HRG39) and 0.12 (ARG10).

The comparison reveals that SM2RAIN-ASCAT dominates the SM2RAIN-CCI across all climate regions. The average percentage improvements in humid, arid and hyper-arid regions were: Bias (27.01%), ubRMSE (19.61%), Theil's U (9.80%), MSEs (24.55%), MSEr (19.41%), and KGE (5.26%) in the humid region; Bias (5.94%), ubRMSE (20.16%), Theil's U (28.20%), MSEs (13.83%), MSEr (29.20%), and KGE (28.12%) in the arid region; and Bias (6.05%), ubRMSE (25.56%), Theil's U (26.83%), MSEs (8.22%), MSEr (24.14%), and KGE (24.72%) in the hyper-arid region.

Figure 4d,e represents the spatial distribution of systematic and random errors of SM2RAIN-ASCAT over four climate regions of Pakistan. Higher random errors were observed in the humid region, which decreases towards the arid and hyper-arid regions. The regional average (median) random errors were 15.41% (13.81%), 57.12% (59.87%), 22.81% (21.97%), and 6.04% (6.75%) in glacial, humid, arid and hyper-arid regions, respectively. Overall the random error average across Pakistan was 25.35%. On the other hand, systematic errors were larger than a random error with an average of 74.65% across the whole of Pakistan. Maximum systematic errors were depicted in hyper-arid (average: 93.96%, median: 93.25%) and glacial (average: 84.76%, median: 86.20%) followed by arid (average: 77.19%, median: 78.03%) and humid (average: 41.59%, median: 40.13%) regions.

Figure 4f displays the spatial distribution of KGE score compared with RG observations. Maximum KGE is observed in the hyper-arid region showing a high performance of SM2RAIN-ASCAT in the region. Smaller KGE scores are observed in the humid and glacial regions depicting poor performance in those mentioned climate regions. However, the KGE score illustrates moderate performance in the arid region. Average (median) KGE scores in glacial, humid, arid and hyper-arid regions were 0.36 (0.34), 0.38 (0.41), 0.64 (0.64), and 0.72 (0.77), respectively.
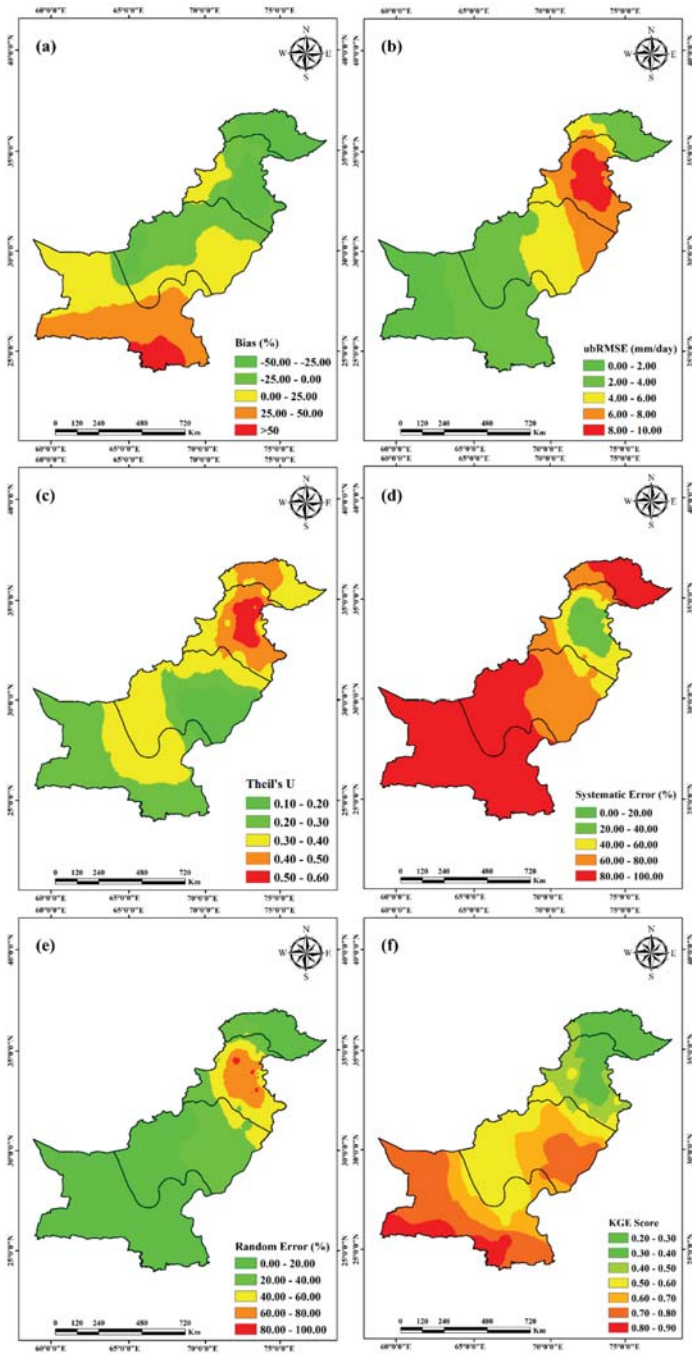
Figure 4. Spatial distribution of bias (**a**), ubRMSE (**b**), Theil's U coefficient (**c**), systematic error (**d**), random error (**e**), and KGE score (**f**) based on a daily scale across Pakistan from SM2RAIN-ASCAT compared to RGs data for the period of 2007–2015.

3.3. Seasonal Evaluation of SM2RAIN-CCI and SM2RAIN-ASCAT

SM2RAIN-CCI and SM2RAIN-ASCAT are compared during four seasons, i.e., pre-monsoon from April to June (A-M-J), monsoon from July to September (J-A-S), post-monsoon from October to November (O-N), and winter from December to March (D-J-F-M) [47,67] for a common period of 2007–2015. Table 4 shows the statistical evaluation of SM2RAIN-CCI in four seasons across Pakistan. The analysis shows poor performance of the precipitation products with an increase in precipitation magnitude and intensity, i.e., monsoon and pre-monsoon seasons. Table 4 shows higher KGE scores in winter season with maximum and minimum values of 0.75 and 0.49, depicting better performance in winter while poor performance in monsoon season (KGE scores ranged between 0.53 (maximum) to 0.19 (minimum)). Similarly, lower biases, ubRMSE, and Theil's U values were observed in winter and post-monsoon seasons. It is concluded from the results that SM2RAIN-CCI shows higher performance in winter followed by post-monsoon season.

Table 4. Seasonal evaluation of SM2RAIN-CCI across Pakistan during 2007-2015.

Season	Statistics	Bias (%)	ubRMSE (mm/Day)	Theil's U	KGE Score
Pre-monsoon (A-M-J)	Maximum	68.21	9.81	0.63	0.59
	Minimum	−36.12	0.73	0.20	0.31
	Average	16.80	4.65	0.45	0.58
	Median	16.19	4.28	0.44	0.56
Monsoon (J-A-S)	Maximum	79.53	19.69	0.68	0.53
	Minimum	−40.44	1.16	0.23	0.19
	Average	20.90	10.52	0.47	0.47
	Median	22.01	10.27	0.48	0.48
Post-monsoon (O-N)	Maximum	56.43	7.45	0.57	0.68
	Minimum	−32.28	0.56	0.18	0.42
	Average	13.76	2.65	0.38	0.65
	Median	14.59	2.12	0.38	0.63
Winter (D-J-F-M)	Maximum	53.88	6.96	0.47	0.75
	Minimum	−33.22	0.39	0.15	0.49
	Average	11.45	3.39	0.34	0.69
	Median	10.98	3.85	0.33	0.67

Table 5 represents the seasonal evaluation of SM2RAIN-ASCAT. SM2RAIN-ASCAT also depicts poor performance in heavy and intense precipitation seasons such as the monsoon and pre-monsoon. Lowest biases and ubRMSE values are observed during the winter season. Theil's U also depicts better performance of SM2RAIN-ASCAT during the winter and post-monsoon seasons. Moreover, similar to SM2RAIN-CCI, the highest KGE scores are observed in winter and post-monsoon seasons while the lowest score in the monsoon season, depicting better performance during low precipitation seasons. Overall, high performances are depicted in winter and post-monsoon seasons. However, SM2RAIN-ASCAT outperformed SM2RAIN-CCI in all four seasons.

3.4. Performance of SM2RAIN-CCI and SM2RAIN-ASCAT in Detecting Precipitation Events

Table 6 shows the daily scale averaged regional categorical metrics of SM2RAIN-CCI and SM2RAIN-ASCAT during a common period of 2007–2015. The categorical metrics are calculated by using a threshold of 1 mm/day. Higher POD (lower FAR) values were detected in the hyper-arid region. The humid region has a lower POD (higher FAR) because of high precipitation magnitude and intensity in the region. Moreover, saturated soil and water storage in hydraulic structures might be the cause of lower POD (higher FAR) in the humid region. Similarly, CSI and the bias score (BS) was also higher in the hyper-arid region and lower in humid region. About the skill of detecting precipitation events, SM2RAIN-ASCAT outperformed SM2RAIN-CCI across three climate regions of Pakistan. In terms of

POD (FAR), SM2RAIN-ASCAT improved the precipitation detection capability by 11.94% (15%) in humid, 10.67% (14.28%) in arid and 8.53% (19.05%) in hyper-arid regions.

Table 5. Seasonal evaluation of SM2RAIN-ASCAT across Pakistan during 2007-2015.

Season	Statistics	Bias (%)	ubRMSE (mm/day)	Theil's U	KGE Score
Pre-monsoon (A-M-J)	Maximum	61.72	8.23	0.54	0.70
	Minimum	−33.20	0.56	0.16	0.42
	Average	11.20	3.97	0.38	0.61
	Median	12.29	4.02	0.40	0.60
Monsoon (J-A-S)	Maximum	71.38	17.94	0.59	0.62
	Minimum	−47.03	0.82	0.20	0.26
	Average	17.82	8.61	0.43	0.55
	Median	17.31	8.47	0.42	0.53
Post-monsoon (O-N)	Maximum	49.21	6.38	0.52	0.78
	Minimum	−29.09	0.52	0.15	0.49
	Average	12.32	2.39	0.36	0.65
	Median	12.62	2.31	0.33	0.66
Winter (D-J-F-M)	Maximum	43.29	5.91	0.44	0.84
	Minimum	−25.90	0.44	0.12	0.55
	Average	8.59	2.04	0.29	0.71
	Median	8.55	1.97	0.28	0.70

Table 6. Daily scale averaged regional categorical metrics of SM2RAIN-CCI (SM2RAIN-ASCAT) during 2007–2015.

Climate Region	Statistics	POD	FAR	CSI	Bias Score
Humid	Average	0.59 (0.67)	0.40 (0.34)	0.57 (0.63)	0.67 (0.73)
	Median	0.62 (0.69)	0.39 (0.35)	0.56 (0.65)	0.72 (0.74)
Arid	Average	0.67 (0.75)	0.28 (0.24)	0.72 (0.79)	0.76 (0.81)
	Median	0.69 (0.73)	0.24 (0.23)	0.71 (0.77)	0.77 (0.84)
Hyper-Arid	Average	0.75 (0.82)	0.21 (0.17)	0.81 (0.90)	0.82 (0.88)
	Median	0.75 (0.81)	0.18 (0.14)	0.82 (0.89)	0.80 (0.89)

*3.5. Evaluation of SM2RAIN-CCI/ASCAT against TMPA*

Previous studies [47,59,67] reported the better performance of TMPA across Pakistan than other PPs, including TRMM-3B42RT, PERSIANN-CDR, PERSIANN-CCS, and CMORPH. Therefore, TMPA was selected to comprehensively assess the performance of SM2RAIN-CCI and SM2RAIN-ASCAT. Three statistical metrics, i.e., Bias, ubRMSE, and Theil's U are used to assess the performance of SM2RAIN-CCI/ASCAT against TMPA during the common period of 2007–2015. Figure 5 shows a comparison of performances of three precipitation products using box-plots. SM2RAIN-CCI data was not available for the glacial region, therefore this region is neglected.

In the humid region, SM2RAIN-CCI underestimated precipitation (negative Biases) and SM2RAIN-ASCAT under/overestimated precipitation (slightly positive Biases), however, TMPA had completely overestimated (positive Biases) precipitation. Moreover, the SM2RAIN-based product showed over/underestimation in the arid region while TMPA overestimated precipitation. A contrasting scenario is observed in the hyper-arid region where TMPA underestimated, and SM2RAIN-CCI/ASCAT overestimated precipitation. Overall, smaller biases are observed when TMPA is compared with RGs, in comparison with SM2RAIN-CCI/ASCAT.

In the case of ubRMSE, TMPA outperformed the SM2RAIN-ASCAT in all three climate regions, while both of them were superior to the SM2RAIN-CCI. The ubRMSE of SM2RAIN-ASCAT and TMPA were almost comparable in humid and hyper-arid regions. Moreover, ubRMSE decreased significantly

from humid region to hyper-arid region, which might be associated with the precipitation magnitude, intensity and any other hydrological (hydraulic) activities in these climate regions.

Among these three PPs, Theil's U coefficient depicts the best performance of TMPA, followed by SM2RAIN-ASCAT and SM2RAIN-CCI. Theil's U shows higher variation (as well as best forecasting accuracy) in the arid region. Overall, TMPA dominated the SM2RAIN-based products across all Pakistan. SM2RAIN-ASCAT and SM2RAIN-CCI need proper bias adjustments before integrating into any hydrological applications.
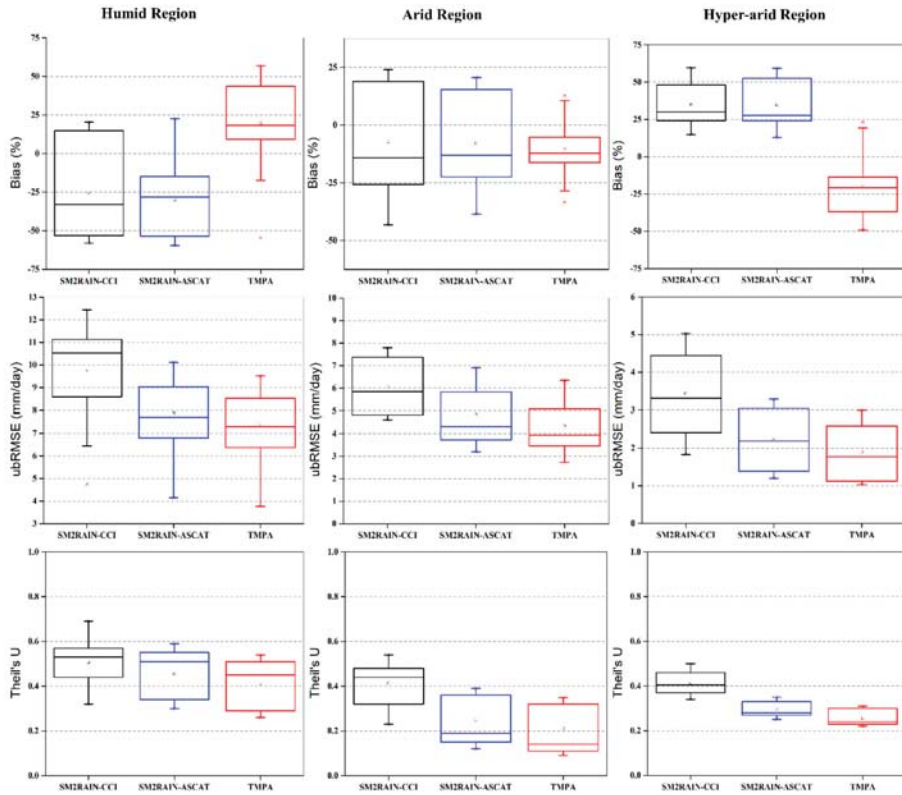
Figure 5. Daily box-plots for Bias (%), ubRMSE (mm/day) and Theil's U for SM2RAIN-CCI (black box) and SM2RAIN-ASCAT (blue box) against TMPA (red box) during the period 2007–2015. The line in each box indicates the median values and each box (top and bottom lines) depicts the 25th and 75th percentile, while the lines perpendicular to boxes (whiskers) extend to the extreme data points.

4. Discussion

The performance of SM2RAIN-CCI (2000–2015) and SM2RAIN-ASCAT (2007–2015) precipitation products against the RGs observations (obtained from PMD and WAPDA departments of Pakistan) was assessed using six statistical and four categorical metrics. Both precipitation products were evaluated over the complex topography and diverse climate of Pakistan, and their performance was also compared with TRMM TMPA 3B42v7 (TMPA) during the period from 2007 to 2015. SM2RAIN-CCI precipitation observations were not available over snow cover, frozen soil, high topographical regions, and rainforests because of the mask used to remove the areas characterized by issues in SM retrieval [45] (shown in Figure 3). Inputs for SM2RAIN-CCI are based on integrating the active and passive ESA CCI SM datasets with bias-corrected and calibrated observations from GPCC-FDD [37]. However,

SM2RAIN-ASCAT observations are obtained from application of the SM2RAIN algorithm to ASCAT SM data without consideration of any filter [32]. The observations are calibrated and bias-corrected based on the ERA5 reanalysis data [43].

Soil moisture-based precipitation products, i.e., SM2RAIN-CCI and SM2RAIN-ASCAT, captured the precipitation both spatially and temporally relatively well. However, both of them tended to fail to estimate the precipitation in the humid region characterized with high precipitation intensity and heavy magnitude (>700 mm), and more specifically in the hydraulically developed region with dams, barrages, headworks, and an extensively developed irrigation canal system (Figures 3 and 4). The performance of SM2RAIN-based products is strongly influenced by spatial and temporal variation of precipitation. Comparatively better performance is observed in the arid and hyper-arid region, which is characterized by low precipitation magnitude (<100 mm). Furthermore, the temporal (seasonal) evaluation also revealed poor performance in monsoon and pre-monsoon seasons, having high precipitation intensity and magnitude. References [9,30] confirmed the poor performance of SM2RAIN algorithm that is unable to adequately estimate precipitation when the soil is close to saturation because the algorithm is unable to derive SM variation when the SM is constant. Similarly, references [41,42] evaluated the performance of SM2RAIN-CCI over north-eastern Brazil and concluded its poor performance in wetter precipitation regimes.

SM2RAIN-CCI and SM2RAIN-ASCAT highly overestimated and underestimated the precipitation in the coastal hyper-arid and mountainous regions of Pakistan, respectively (Figures 3a and 4a). This can be considered as SM retrieval uncertainties, which significantly deteriorate the accuracy of precipitation estimates from the SM2RAIN algorithm (i.e., due to error propagation) [43]. Underestimation of precipitation is associated to poor infiltration capacity and soil moisture storage capacities of rocks (hills), permanent snow and glacier cover, hydraulic developmental activities, and intense and heavy precipitation during monsoon and pre-monsoon seasons, while overestimation is associated to low precipitation intensity and magnitude.

Though SM2RAIN-ASCAT underestimated the precipitation significantly in humid and glacial regions, the locations of heavy rainfall were relatively accurately detected as compared to SM2RAIN-CCI (Figures 3 and 4). Larger biases and uncertainties at those locations are due to the impact of complex topography and diverse climate, which have been well documented in previous literature [47,59,67,68]. Higher positive or negative biases for SM2RAIN-based products are expected over densely vegetated and forest regions (glacial and hilly areas of the humid region) where signals of satellite sensors do not breach into dense vegetation [69], or in the region where the soil remains saturated for longer period of time such as flooded regions (the humid region, especially the hydraulically developed areas) [9], respectively. Considering other statistical metrics, all the metrics showed poor performance in precipitation dominated regions and seasons (Figures 3 and 4, Tables 4 and 5). References [9,41] confirm the findings of the current study.

Concerning the error components, the analysis revealed that SM2RAIN-CCI and SM2RAIN-ASCAT were similar in the spatial distribution of systematic and random errors (Figure 3d,e and Figure 4d,e). The results revealed that systematic error contributed significantly in comparison with random errors, supporting the findings of previous studies conducted by Prakash [6] and Paredes-Trejo, Barbosa and dos Santos [41] in India and Brazil, respectively. Therefore, these products require refinement and corrections before their integration into a particular hydrological application [70].

Seasonal evaluation of SM2RAIN-CCI and SM2RAIN-ASCAT presented a better performance in moderate to low precipitation seasons, i.e., post-monsoon and winter seasons (Tables 4 and 5). During these two seasons, precipitation has had enough time to infiltrate into the soil and then saturate; therefore, the precipitation estimation capabilities of SM2RAIN-based precipitation products are high and hence produce relatively accurate results. However, precipitation in summer (monsoon season), which is distinguished by intense and short duration precipitation events, is the influential factor that affected the performance of both SM2RAIN-based products in terms of biases. Therefore, the possibility

exists that the sensors might have not accurately estimated (or missed) the precipitation during the period [41].

The TMPA precipitation product performed reasonably well across different climate regions when compared with SM2RAIN-CCI and SM2RAIN-ASCAT (Figure 5). A different trend is observed in the humid and arid regions where SM2RAIN-based products overestimate and underestimates the precipitation; TMPA overestimates the precipitation. In contrast, in the hyper-arid region, TMPA underestimates while SM2RAIN-based product overestimates the precipitation. SM2RAIN-CCI depicted poor performance, implying high uncertainties when compared to SM2RAIN-ASCAT and TMPA across all three climate regions. TMPA showed comparatively good agreement with gauged based observations in different climate regions and different seasons relative to the SM2RAIN-based products.

The performance of SM2RAIN-based products across glacial, humid, arid and hyper-arid regions during the common period (2007-2015) are summarized in Tables 7 and 8.

Table 7. Summary of the performance of SM2RAIN-CCI datasets across the four climate regions of Pakistan during 2007–2015.

Climate Region	Evaluation Metrics	Performance	Reasons
Humid	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	*Underestimation (majority) **Poor performance ***Lower magnitude **** Higher magnitude	*High underestimation due to hydraulically developed area ** Due to intense and heavy precipitation
Arid	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	*Underestimation over mountainous areas while overestimation in plain areas **Comparatively better ubRMSE while higher accuracy is detected over agricultural areas and low accuracy at over mountainous areas **High KGE in plain areas while low in mountain areas ***Highest in the south while higher in rest of the region ****Lowest in the south while lower in rest of the region	*Underestimation is due to poor infiltration capacity of rock in the east of the region, moreover, snow cover in winter also contributes to underestimation **Higher performance in the region in terms of ubRMSE and Theil's U is due to agriculture and moderate precipitation in the region
Hyper-arid	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	*Overestimation in all the regions (higher in south-east) **Lowest ubRMSE while moderate to lower forecasting accuracy **Moderate KGE ***Highest among all climate regions ****Lowest among all climate regions	*Overestimation due to seawater intrusion and minimum precipitation in the region

Table 8. Summary of the performance of SM2RAIN-ASCAT dataset across four climate regions of Pakistan during 2007–2015.

Climate Region	Evaluation Metrics	Performance	Reasons
Glacial	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	*Underestimation, high in east while relatively low in west **Relatively better performance in east **Poor forecasting in middle of the region **Lower KGE ***Highest toward the east ****Lowest toward the east	*Underestimation due to permanent snow and glacier cover **Errors due to low infiltration and storage capacity in the region
Humid	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	*Underestimation in the north-east while overestimation in the west of the region **Poor performance **Poor forecasting accuracy **Lower KGE score ***Lower systematic error ****High random error	*High underestimation is due to heavy and intense precipitation in north-east *Hydraulically developed areas and flooded regions
Arid	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	*Underestimation in the mountainous region while overestimation in plain agricultural lands **Moderate RMSE and high forecasting accuracy in plain areas **High KGE in plain areas while low over mountain areas ***Highest over elevated areas while moderate in plain areas ****Lowest over mountain areas while moderate in plain areas	*Underestimation in the west due to snow cover in winter season, poor infiltration and soil water storage capacity of hilly areas **Moderate precipitation and agricultural activities which preserve soil moisture
Hyper-arid	*Bias **ubRMSE **Theil's U **KGE ***Systematic error ****Random error	**Overestimation in all the regions (higher in south-east) **Lowest ubRMSE (west) while moderate (extreme south along the coast of Arabian Sea) to lower forecasting (north of the region) accuracy **High KGE along the coast, moderate to relatively low in rest of the region ***Highest systematic error ****Lowest random error	*Hyper-arid region receives very low precipitation. Seas water intrusion might be the reason for the overestimation **Low precipitation magnitude in the region causes lower uncertainties

Based on spatial and temporal evaluation of SM2RAIN-CCI/ASCAT precipitation products across Pakistan along different climate regions, and the obtained results, the following recommendations are suggested for further studies and applications: (1) Agricultural water management and irrigation scheduling in the arid region (the Punjab province, which is agricultural hub of the country), where SM2RAIN-based PPs performed comparatively better, (2) An early flood warning system (EFWS) and flood simulation, where soil moisture (besides the precipitation intensity) plays important role, and bias-corrected SM2RAIN-based products will be helpful for reducing the impact of flood on society, (3) Drought monitoring in drought-prone arid and hyper-arid regions, including hydrological drought (shortage in water storage as well as net precipitation at same time) and meteorological drought (shortage in the catchment's water fluxes such as precipitation), (4) Vegetation and crop growth monitoring, and (5) Groundwater modeling and rainwater harvesting studies.

5. Conclusions

In this study, the performance of SM2RAIN-CCI and SM2RAIN-ASCAT were assessed extensively across Pakistan using the rain gauge (RGs) observations. The assessment was carried across four different climate regions of Pakistan, namely glacial, humid, arid and hyper-arid regions during the period 2000–2015 (for SM2RAIN-CCI) and 2007–2015 for (SM2RAIN-ASCAT) on a daily scale. Three statistical metrics, i.e., Bias (%), ubRMSE (mm/day), and Theil's U coefficient, and four categorical metrics, i.e., POD, FAR, CSI and Bias score, were used to assess the performance of SM2RAIN-based products. Moreover, the total mean square error was also decomposed to random and systematic error components, while KGE was also used to evaluate the performance of PPs. Finally, SM2RAIN-CCI and SM2RAIN-ASCAT were compared and evaluated against an extensively evaluated satellite precipitation product, i.e., TRMM TMPA 3B42 v7 (TMPA). Based on the results, following conclusions were drawn:

(1) Both SM2RAIN-CCI and SM2RAIN-ASCAT underestimated precipitation in north-east of the humid region and hydraulically developed areas, i.e., regions with barrages, headworks, dams and an extensively developed irrigation canal system. Maximum underestimation for SM2RAIN-CCI (SM2RIAN-ASCAT) was 58.04% (42.36%).

(2) SM2RAIN-based product underestimated precipitation over mountainous and glacial regions (SM2RAIN-ASCAT). In mountainous regions, the maximum underestimation of SM2RAIN-CCI was 43.16% (at ARG12) and 34.60% (at GRG17). This underestimation is due to low infiltration and soil water storage capacities of these mountainous areas.

(3) Besides, overestimation is also observed in coastal hyper-arid areas (near the coast of Arabian Sea), which can be considered due to soil moisture retrieval errors, i.e., propagation errors that affect the quality of precipitation estimates from the SM2TAIN algorithm and soil water intrusion from the Arabian Sea. Maximum overestimation for SM2RAIN-CCI (SM2RAIN-ASCAT) was 59.59% (52.35%) at HARG20.

(4) Based on the Theil's U, high forecasting accuracies for SM2RAIN-based products were observed in the arid region (plain areas). Theil's U for SM2RAIN-CCI (SM2RAIN-ASCAT) was 0.23 at ARG15 (0.15 at ARG16).

(5) SM2RAIN-ASCAT outperformed SM2RAIN-CCI across all climate regions of Pakistan. The average percentage improvements were: Bias (27.01% in humid, 5.94% in arid, and 6.05% in hyper-arid), ubRMSE (19.61% in humid, 20.16% in arid, and 25.56% in hyper-arid), Theil's U (9.80% in humid, 28.80% in arid, and 26.83% in hyper-arid), MSEs (24.55% in humid, 13.83% in arid, and 8.22% in hyper-arid), MSEr (19.41% in humid, 29.20% in arid, and 24.14% in hyper-arid) and KGE score (5.26% in humid, 28.12% in arid, and 24.72% in hyper-arid).

(6) The reliability of SM2RAIN-based products is dependent on topography and climate (precipitation variability), for instance, SM2RAIN-based products failed to estimate precipitation in mountains and semi-arid climate (west of arid region).

(7) Among the systematic and random errors of SM2RAIN-CCI and SM2RAIN-ASCAT, systematic error dominated random error across all Pakistan, which suggest that the products need bias correction before integrating them in any hydrological application.

(8) SM2RAIN-based products failed to accurately capture the intense (short duration) and heavy (long term events) precipitation during the monsoon and pre-monsoon seasons. Average values of statistical metrics during the monsoon season for SM2RAIN-CCI (SM2RAIN-ASCAT) were 20.90% (17.82%), 10.52 mm/day (8.61 mm/day), 0.47 (0.43), and 0.47 (0.55) for Bias, ubRMSE, Theil's U, and KGE score, respectively. However, the performance improved from post-monsoon to the winter season as the precipitation magnitude decreased. Average values for statistical metrics during winter seasons for SM2RAIN-CCI (SM2RAIN-ASCAT) were 11.45% (8.59%), 3.39 mm/day (2.04 mm/day), 0.34 (0.29), and 0.69 (0.71) for Bias, ubRMSE, Theil's U, and KGE score, respectively.

(9)   Comparison of SM2RAIN-based products with TMPA revealed the better performance of TMPA across all climate regions in terms of all statistical metrics.

References

1.   Maggioni, V.; Massari, C. On the performance of satellite precipitation products in riverine flood modeling: A review. *J. Hydrol.* **2018**, *558*, 214–224. [CrossRef]

2.   Skofronick-Jackson, G.; Petersen, W.A.; Berg, W.; Kidd, C.; Stocker, E.F.; Kirschbaum, D.B.; Kakar, R.; Braun, S.A.; Huffman, G.J.; Iguchi, T. The global precipitation measurement (GPM) mission for science and society. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1679–1695. [CrossRef] [PubMed]

3.   Brunetti, M.; Melillo, M.; Peruccacci, S.; Ciabatta, L.; Brocca, L. How far are we from the use of satellite rainfall products in landslide forecasting? *Remote Sens. Environ.* **2018**, *210*, 65–75. [CrossRef]

4.   Forootan, E.; Khaki, M.; Schumacher, M.; Wulfmeyer, V.; Mehrnegar, N.; van Dijk, A.; Brocca, L.; Farzaneh, S.; Akinluyi, F.; Ramillien, G. Understanding the global hydrological droughts of 2003–2016 and their relationships with teleconnections. *Sci. Total Environ.* **2019**, *650*, 2587–2604. [CrossRef] [PubMed]

5.   Pendergrass, A.G.; Knutti, R. The uneven nature of daily precipitation and its change. *Geophys. Res. Lett.* **2018**, *45*, 11,980–911,988. [CrossRef]

6.   Prakash, S. Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *J. Hydrol.* **2019**, *571*, 50–59. [CrossRef]

7.   Zambrano, F.; Wardlow, B.; Tadesse, T.; Lillo-Saavedra, M.; Lagos, O. Evaluating satellite-derived long-term historical precipitation datasets for drought monitoring in Chile. *Atmos. Res.* **2017**, *186*, 26–42. [CrossRef]

8.   Rinaldo, A.; Bertuzzo, E.; Mari, L.; Righetto, L.; Blokesch, M.; Gatto, M.; Casagrandi, R.; Murray, M.; Vesenbeckh, S.M.; Rodriguez-Iturbe, I. Reassessment of the 2010–2011 Haiti cholera outbreak and rainfall-driven multiseason projections. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6602–6607. [CrossRef]

9.   Brocca, L.; Ciabatta, L.; Massari, C.; Moramarco, T.; Hahn, S.; Hasenauer, S.; Kidd, R.; Dorigo, W.; Wagner, W.; Levizzani, V. Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *J. Geophys. Res. Atmos.* **2014**, *119*, 5128–5141. [CrossRef]

10. 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]

11. Villarini, G.; Mandapaka, P.V.; Krajewski, W.F.; Moore, R.J. Rainfall and sampling uncertainties: A rain gauge perspective. *J. Geophys. Res. Atmos.* **2008**, *113*. [CrossRef]

12. Kidd, C.; Becker, A.; Huffman, G.J.; Muller, C.L.; Joe, P.; Skofronick-Jackson, G.; Kirschbaum, D.B. So, how much of the Earth's surface is covered by rain gauges? *Bull. Am. Meteorol. Soc.* **2017**, *98*, 69–78. [CrossRef] [PubMed]

13. Ebert, E.E.; Janowiak, J.E.; Kidd, C. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 47–64. [CrossRef]

14. Massari, C.; Crow, W.; Brocca, L. An assessment of the performance of global rainfall estimates without ground-based observations. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 4347–4361. [CrossRef]

15. 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]

16. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116. [CrossRef]

17. Abera, W.; Brocca, L.; Rigon, R. Comparative evaluation of different satellite rainfall estimation products and bias correction in the Upper Blue Nile (UBN) basin. *Atmos. Res.* **2016**, *178*, 471–483. [CrossRef]

18. Kimani, M.; Hoedjes, J.; Su, Z. An assessment of satellite-derived rainfall products relative to ground observations over East Africa. *Remote Sens.* **2017**, *9*, 430. [CrossRef]

19. 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]

20. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 150066. [CrossRef]

21. Joyce, R.J.; Xie, P. Kalman filter–based CMORPH. *J. Hydrometeorol.* **2011**, *12*, 1547–1563. [CrossRef]

22. Sorooshian, S.; Hsu, K.-L.; Gao, X.; Gupta, H.V.; Imam, B.; Braithwaite, D. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2035–2046. [CrossRef]

23. Nguyen, P.; Ombadi, M.; Sorooshian, S.; Hsu, K.; AghaKouchak, A.; Braithwaite, D.; Ashouri, H.; Thorstensen, A.R. The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 5801–5816. [CrossRef]

24. Ushio, T.; Sasashige, K.; Kubota, T.; Shige, S.; Okamoto, K.I.; Aonashi, K.; Inoue, T.; Takahashi, N.; Iguchi, T.; Kachi, M. A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteorol. Soc. Jpn. Ser. II* **2009**, *87*, 137–151. [CrossRef]

25. Crow, W.; van Den Berg, M.; Huffman, G.; Pellarin, T. Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). *Water Resour. Res.* **2011**, *47*. [CrossRef]

26. Pellarin, T.; Louvet, S.; Gruhier, C.; Quantin, G.; Legout, C. A simple and effective method for correcting soil moisture and precipitation estimates using AMSR-E measurements. *Remote Sens. Environ.* **2013**, *136*, 28–36. [CrossRef]

27. Brocca, L.; Moramarco, T.; Melone, F.; Wagner, W. A new method for rainfall estimation through soil moisture observations. *Geophys. Res. Lett.* **2013**, *40*, 853–858. [CrossRef]

28. Koster, R.D.; Brocca, L.; Crow, W.T.; Burgin, M.S.; De Lannoy, G.J. Precipitation estimation using L-band and C-band soil moisture retrievals. *Water Resour. Res.* **2016**, *52*, 7213–7225. [CrossRef]

29. Tarpanelli, A.; Massari, C.; Ciabatta, L.; Filippucci, P.; Amarnath, G.; Brocca, L. Exploiting a constellation of satellite soil moisture sensors for accurate rainfall estimation. *Adv. Water Resour.* **2017**, *108*, 249–255. [CrossRef]

30. Ciabatta, L.; Brocca, L.; Massari, C.; Moramarco, T.; Puca, S.; Rinollo, A.; Gabellani, S.; Wagner, W. Integration of satellite soil moisture and rainfall observations over the Italian territory. *J. Hydrometeorol.* **2015**, *16*, 1341–1355. [CrossRef]

31. Ciabatta, L.; Massari, C.; Brocca, L.; Reimer, C.; Hann, S.; Dorigo, W.; Wagner, W. Using Python®Language for the Validation of the CCI Soil Moisture Products Via SM2RAIN. *PeerJ Preprints* **2016**, *4*, e2131v4. [CrossRef]

32. Brocca, L.; Massari, C.; Ciabatta, L.; Moramarco, T.; Penna, D.; Zuecco, G.; Pianezzola, L.; Borga, M.; Matgen, P.; Martínez-Fernández, J. Rainfall estimation from in situ soil moisture observations at several sites in Europe: An evaluation of the SM2RAIN algorithm. *J. Hydrol. Hydromech.* **2015**, *63*, 201–209. [CrossRef]

33. Chiaravalloti, F.; Brocca, L.; Procopio, A.; Massari, C.; Gabriele, S. Assessment of GPM and SM2RAIN-ASCAT rainfall products over complex terrain in southern Italy. *Atmos. Res.* **2018**, *206*, 64–74. [CrossRef]

34. Massari, C.; Camici, S.; Ciabatta, L.; Brocca, L. Exploiting satellite-based surface soil moisture for flood forecasting in the Mediterranean area: State update versus rainfall correction. *Remote Sens.* **2018**, *10*, 292. [CrossRef]

35. Brocca, L.; Pellarin, T.; Crow, W.T.; Ciabatta, L.; Massari, C.; Ryu, D.; Su, C.H.; Rüdiger, C.; Kerr, Y. Rainfall estimation by inverting SMOS soil moisture estimates: A comparison of different methods over Australia. *J. Geophys. Res. Atmos.* **2016**, *121*, 12,062–12,079. [CrossRef]

36. Ciabatta, L.; Marra, A.C.; Panegrossi, G.; Casella, D.; Sanò, P.; Dietrich, S.; Massari, C.; Brocca, L. Daily precipitation estimation through different microwave sensors: Verification study over Italy. *J. Hydrol.* **2017**, *545*, 436–450. [CrossRef]

37. Ciabatta, L.; Massari, C.; Brocca, L.; Gruber, A.; Reimer, C.; Hahn, S.; Paulik, C.; Dorigo, W.; Kidd, R.; Wagner, W. SM2RAIN-CCI: A new global long-term rainfall data set derived from ESA CCI soil moisture. *Earth Syst. Sci. Data* **2018**, *10*, 267. [CrossRef]

38. Wanders, N.; Pan, M.; Wood, E.F. Correction of real-time satellite precipitation with multi-sensor satellite observations of land surface variables. *Remote Sens. Environ.* **2015**, *160*, 206–221. [CrossRef]

39. Massari, C.; Brocca, L.; Moramarco, T.; Tramblay, Y.; Lescot, J.-F.D. Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall. *Adv. Water Resour.* **2014**, *74*, 44–53. [CrossRef]

40. Román-Cascón, C.; Pellarin, T.; Gibon, F.; Brocca, L.; Cosme, E.; Crow, W.; Fernández-Prieto, D.; Kerr, Y.H.; Massari, C. Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX. *Remote Sens. Environ.* **2017**, *200*, 295–310. [CrossRef]

41. Paredes-Trejo, F.; Barbosa, H.; dos Santos, C.A. Evaluation of the Performance of SM2RAIN-Derived Rainfall Products over Brazil. *Remote Sens.* **2019**, *11*, 1113. [CrossRef]

42. Paredes-Trejo, F.; Barbosa, H.; Rossato Spatafora, L. Assessment of SM2RAIN-Derived and State-of-the-Art Satellite Rainfall Products over Northeastern Brazil. *Remote Sens.* **2018**, *10*, 1093. [CrossRef]

43. Brocca, L.; Crow, W.T.; Ciabatta, L.; Massari, C.; De Rosnay, P.; Enenkel, M.; Hahn, S.; Amarnath, G.; Camici, S.; Tarpanelli, A. A review of the applications of ASCAT soil moisture products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2285–2306. [CrossRef]

44. Crow, W.T.; Huffman, G.J.; Bindlish, R.; Jackson, T.J. Improving satellite-based rainfall accumulation estimates using spaceborne surface soil moisture retrievals. *J. Hydrometeorol.* **2009**, *10*, 199–212. [CrossRef]

45. Camici, S.; Ciabatta, L.; Massari, C.; Brocca, L. How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean area. *J. Hydrol.* **2018**, *563*, 950–961. [CrossRef]

46. Brocca, L.; Filippucci, P.; Hahn, S.; Ciabatta, L.; Massari, C.; Camici, S.; Schüller, L.; Bojkov, B.; Wagner; Wolfgang, W. SM2RAIN-ASCAT (2007–2018): Global daily satellite rainfall from ASCAT soil moisture. *Earth Syst. Sci. Data* **2019**. Under Review. [CrossRef]

47. Rahman, K.U.; Shang, S.; Shahid, M.; Li, J. Developing an Ensemble Precipitation Algorithm from Satellite Products and Its Topographical and Seasonal Evaluations over Pakistan. *Remote Sens.* **2018**, *10*, 1835. [CrossRef]

48. Balkhair, K.S.; Rahman, K.U. Sustainable and economical small-scale and low-head hydropower generation: A promising alternative potential solution for energy generation at local and regional scale. *Appl. Energy* **2017**, *188*, 378–391. [CrossRef]

49. Salma, S.; Shah, M.; Rehman, S. Rainfall trends in different climate zones of Pakistan. *Pak. J. Meteorol.* **2012**, *9*, 37–47.

50. Dimri, A.; Niyogi, D.; Barros, A.; Ridley, J.; Mohanty, U.; Yasunari, T.; Sikka, D. Western disturbances: A review. *Rev. Geophys.* **2015**, *53*, 225–246. [CrossRef]

51. Asmat, U.; Athar, H. Run-based multi-model interannual variability assessment of precipitation and temperature over Pakistan using two IPCC AR4-based AOGCMs. *Theor. Appl. Climatol.* **2017**, *127*, 16. [CrossRef]

52. Liu, Y.Y.; Parinussa, R.; Dorigo, W.A.; De Jeu, R.A.; Wagner, W.; Van Dijk, A.; McCabe, M.F.; Evans, J. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 425–436. [CrossRef]

53. Dorigo, W.; Gruber, A.; De Jeu, R.; Wagner, W.; Stacke, T.; Loew, A.; Albergel, C.; Brocca, L.; Chung, D.; Parinussa, R. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sens. Environ.* **2015**, *162*, 380–395. [CrossRef]

54. Cui, C.; Xu, J.; Zeng, J.; Chen, K.-S.; Bai, X.; Lu, H.; Chen, Q.; Zhao, T. Soil moisture mapping from satellites: An intercomparison of SMAP, SMOS, FY3B, AMSR2, and ESA CCI over two dense network regions at different spatial scales. *Remote Sens.* **2018**, *10*, 33. [CrossRef]

55. Wagner, W.; Hahn, S.; Kidd, R.; Melzer, T.; Bartalis, Z.; Hasenauer, S.; Figa-Saldaña, J.; de Rosnay, P.; Jann, A.; Schneider, S. The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorol. Z.* **2013**, *22*, 5–33. [CrossRef]

56. Dorigo, W.; Wagner, W.; Albergel, C.; Albrecht, F.; Balsamo, G.; Brocca, L.; Chung, D.; Ertl, M.; Forkel, M.; Gruber, A. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.* **2017**, *203*, 185–215. [CrossRef]

57. Melo, D.C.D.; Xavier, A.C.; Bianchi, T.; Oliveira, P.T.; Scanlon, B.R.; Lucas, M.C.; Wendland, E. Performance evaluation of rainfall estimates by TRMM Multi-satellite Precipitation Analysis 3B42V6 and V7 over Brazil. *J. Geophys. Res. Atmos.* **2015**, *120*, 9426–9436. [CrossRef]

58. Caparoci Nogueira, S.; Moreira, M.; Lordelo Volpato, M. Evaluating precipitation estimates from Eta, TRMM and CHRIPS Data in the south-southeast region of Minas Gerais State—Brazil. *Remote Sens.* **2018**, *10*, 313. [CrossRef]

59. Iqbal, M.F.; Athar, H. Validation of satellite based precipitation over diverse topography of Pakistan. *Atmos. Res.* **2018**, *201*, 247–260. [CrossRef]

60. Bliemel, F. *Theil's Forecast Accuracy Coefficient: A Clarification*; SAGE Publications Sage CA: Los Angeles, CA, USA, 1973.

61. AghaKouchak, A.; Mehran, A.; Norouzi, H.; Behrangi, A. Systematic and random error components in satellite precipitation data sets. *Geophys. Res. Lett.* **2012**, *39*. [CrossRef]

62. Huang, A.; Zhao, Y.; Zhou, Y.; Yang, B.; Zhang, L.; Dong, X.; Fang, D.; Wu, Y. Evaluation of multisatellite precipitation products by use of ground-based data over China. *J. Geophys. Res. Atmos.* **2016**, *121*, 10654–10675. [CrossRef]

63. Kling, H.; Fuchs, M.; Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* **2012**, *424*, 264–277. [CrossRef]

64. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]

65. Toté, C.; Patricio, D.; Boogaard, H.; Van Der Wijngaart, R.; Tarnavsky, E.; Funk, C. Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. *Remote Sens.* **2015**, *7*, 1758–1776. [CrossRef]

66. Zia, I.; Zafar, H.; Shahzad, M.I.; Meraj, M.; Kazmi, J.H. Assessment of sea water inundation along Daboo Creek area in Indus Delta Region, Pakistan. *J. Ocean. Univ. China* **2017**, *16*, 1055–1060. [CrossRef]

67. 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]

68. 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]

69. Rodríguez-Fernández, N.J.; Muñoz Sabater, J.; Richaume, P.; Rosnay, P.D.; Kerr, Y.H.; Albergel, C.; Drusch, M.; Mecklenburg, S. SMOS near-real-time soil moisture product: Processor overview and first validation results. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5201–5216. [CrossRef]

70. Bhatti, H.; Rientjes, T.; Haile, A.; Habib, E.; Verhoef, W. Evaluation of bias correction method for satellite-based rainfall data. *Sensors* **2016**, *16*, 884. [CrossRef]

Performance of the State-Of-The-Art Gridded Precipitation Products over Mountainous Terrain: A Regional Study over Austria

Ehsan Sharifi [1,2,*], Josef Eitzinger [1] and Wouter Dorigo [3]

[1] Institute of Meteorology and Climatology, University of Natural Resources and Life Sciences (BOKU), 1180 Vienna, Austria
[2] Department of Meteorology and Geophysics, University of Vienna, 1090 Vienna, Austria
[3] Department of Geodesy and Geo-Information, Vienna University of Technology (TU-Wien), 1040 Vienna, Austria
*   Correspondence: ehsan.sharifi@boku.ac.at or ehsan.sharifi@univie.ac.at; Tel.: +43-676-976-6304

Abstract: During the last decade, satellite-based precipitation products have been believed to be a potential source for forcing inputs in hydro-meteorological and agricultural models, which are essential especially over the mountains area or in basins where ground gauges are generally sparse or nonexistent. This study comprehensively evaluates several newly released precipitation products, i.e., MSWEP-V2.2, IMERG-V05B, IMERG-V06A, IMERG-V05-RT, ERA5, and SM2RAIN-ASCAT, at daily and monthly time-scales over Austria. We show that all the examined products are able to reproduce the spatial precipitation distribution over the country. MSWEP, followed by IMERG-V05B and -V06A, show the strongest agreement with in situ observations and perform better than other products with respect to spatial patterns and statistical metrics. Both IMERG and ERA5 products seem to have systematic precipitation overestimation at the monthly time-scale. IMERG-V06A performs slightly better than IMERG-V05B. With respect to heavy precipitation (P > 10 mm/day), MSWEP compare to other products demonstrate advantages in detecting precipitation events with a higher spatial average of probability of detection (POD) and lower false alarm ratio (FAR) scores skill (0.74 and 0.28), while SM2RAIN and ERA5 reveal lower POD (0.35) and higher FAR (0.56) in this precipitation range in comparison with other products. However, the ERA5 and MSWEP indicate robust average POD and FAR values with respect to light/moderate precipitation (10 mm > P ≥ 0.1 mm) with 0.94 and 0.11, respectively. Such robustness of MSWEP may be rooted in applying the daily rain gauges in calibration processes. Moreover, although all products accurately map the spatial precipitation distribution they still have difficulty capturing the effects of topography on precipitation. The overall performance of the precipitation products was lower in the peripheries of the study area where most stations are situated in the mountainous area and was higher over the low altitude regions. However, according to our analysis of the considered products, MSWEP-V2.2, followed by IMERG-V06S and -V05B, are the most suitable for driving hydro-meteorological, agricultural, and other models over mountainous terrain.

Keywords: satellite-based precipitation; elevation; extreme events; IMERG-V05B and V06A; MSWEP; ERA5; SM2RAIN

1. Introduction

Droughts and floods are water-related natural phenomena which have large negative impacts on society and activities related to agriculture, and local economies. Drought is one of the most important natural disasters, since it affects wide areas for long time (months to years) and, thus, has a serious impact on regional or countries economic performance, etc.

In recent decades, large-scale extreme events (i.e., droughts) have been observed in many places around the world leading to high negative impacts on economic, ecological resources, food shortages, etc. However, floods are among the most destructive natural phenomena, declaring more lives and leading to more property damage than any other natural events. The reliable and accurate drought and flood information have been more interested for a variety of authorities, such as water managers, policy makers, researchers, farmers, etc., for effective management [1].

Since precipitation is the most important factor of the aforementioned phenomena, knowing the locations, domain, and length of precipitation is essential to understand, predict, and mitigate the impact of such disasters. Irrespective of the less accessible mountainous and oceanic regions, compared to ground-based measurements, such as gauges and radars, satellite precipitation estimates (SPE) products are able to cover the precipitation system at a nearly global-scale. Generally, in situ observations are often subject to wind effects, many missing values, insufficient number of stations, or sparse gauge networks, particularly in mountainous or desert areas [2,3]. Moreover, gridded daily surface precipitation data are important for many water-related applications, such as drought and flood monitoring systems. Rapid growth in computer technology and the remote sensing area help observations processed from satellites, individually, and merge them with other data sources to provide a better understanding of the precipitation spatial visualization. The information derived from SPEs provides tremendous potential for identification, monitoring, and assessment of droughts, flood, etc., especially for regions with sparse rain gauges or limited radar coverage [4].

However, the precision of SPEs at spatiotemporal representations has a great influence on the effective predictions of natural hazard, climate impacts, etc.; therefore, accuracy analysis of the new precipitation products is often applied before it can be employed in decision-making activities [5]. The satellite/gridded data produce the area average of precipitation in contrast with point measurements obtained with rain gauges. Earlier studies have shown that diverse altitude and geographic and climate conditions have greatly impact on the accuracy and performance of satellite or other precipitation products [2,5–11]. For instance, in a study by Gottschalck et al. [12] the overestimation of 3B42-RT over the Central United States is attributed to misclassification of cold cirrus clouds as precipitating systems. In another study, Dinku et al. [13] demonstrated that topography plays a significant role in SPE due to the weakness of algorithms to detect orographically-induced precipitation. Numerous studies have shown that an appropriate interpolation method might develop a gridded dataset using the rain gauges, but the obtained dataset is dependent on both adequate underlying station observations and the use of an appropriate interpolation technique to produce high-resolution gridded point estimates prior to the creation of area-averages grid values [14]. Thaler et al. [15] used different gridded precipitation data in an agronomic application and analyzed how different products influence a crop model application. The Austrian radar network (Austrocontrol) is operated consisting of four radar stations distributed across the country. However, due to the mountainous area and terrestrial characteristics of the country, in many western regions of Austria radar data have limitation to use particularly during wintertime, when precipitation may originate from rather shallow cloud systems.

In recent years, with the rapid development of remote sensing technique, more and more quasi-global satellite precipitation products have been produced and released to the public. The objective of this paper is inter-comparison of the recently released gridded precipitation products (e.g., MSWEP-V2.2, IMERG-V05B, IMERG-V05-RT -V06A, ERA5, and SM2RAIN) against in situ observations. The question motivating this study is "To what extent have the recently released gridded precipitation products, improved as compared to a dense rain-gauge network?" The total error is broken down into estimation and detection of precipitation in order to assess the algorithm performance than can highlight the weakness and strength of those algorithms and assist the developer to improve those aspects that have greater need.

In this study, the gauge-based measurements are directly compared against the corresponding gridded data, enabling us to identify the best precipitation estimated product and how close they are to source data (stations).

It is noteworthy to mention that, previously, Sharifi et al. [3] conducted a study to evaluate the reliability of the IMERG-V04 final-run (FR) and real-time (RT) products against the Central Institute for Meteorology and Geodynamics (ZAMG) stations over Northeast Austria. However, an evaluation of the performances of ERA5, MSWEP-V2.2, IMERG-V5B, IMERG-V6, IMERG-V5-RT, and SM2RAIN products has not been conducted over Austria. Therefore, the aforementioned products and a highly dense in situ precipitation network provided by the Federal Ministry for Sustainability and Tourism (BMNT)-Austria (882 stations) are selected in this study. Lastly, this study will be useful since it will provide the reference for precipitation monitoring and regional climate prediction across Austria.

2. Data and Study Area

The data used in this study are described below and its summaries and characteristics have been shown in Table 1.

Table 1. Characteristics of the precipitation products.

Products	Temporal Resolution	Spatial Resolution	Coverage	Availability
IMERG-V05B	30 min	0.1°	60°N–60°S	March 2014–present
IMERG-V05-RT	30 min	0.1°	60°N–60°S	March 2014–present
IMERG-V06A	30 min	0.1°	60°N–60°S	June 2014–present
MSWEP-V2.2	3 h	0.1°	90°N–90°S	1979–2017
ERA5	1 h	0.28°	90°N–90°S	1950–present
SM2RAIN	Daily	12.5 km	90°N–90°S	2007–2018

2.1. In Situ Data

In this study, we use precipitation data from 782 stations provided by BMNT, followed by proper quality-control checks within Austria, in the range of daily time-scale between 1985 and 2016.

However, there are other stations provided by ZAMG-Austria distributed throughout the country. In this study the in situ observation network provided by BMNT (hereafter: eHYD) is used for two reasons: first, the eHYD network is much denser than the ZAMG stations. Second, eHYD provides gauge observations that are entirely independent of the gridded precipitation products that are assessed in this study, while a large number of ZAMG observations are used by the Global Precipitation Climatology Centre (GPCC) products, which, in turn, are integrated in several of the products assessed in this study (i.e., IMERG). As an example, Sharifi et al. [3] reported that over the northeast of the country 51 out of 62 ZAMG synoptic stations are shared with GPCC, which contains a major part of the total ZAMG gauges (82%). Since the satellite products used in this study are used with GPCC gauge data for calibration, we utilize the independent eHYD observations in order to avoid any misleading results. Spatial distribution of both eHYD and ZAMG stations and elevation map of Austria are shown in Figure 1.

It is worthy to mention that stratiform, convective, and orographic precipitation types occur over the country [16].

2.2. Precipitation Products

2.2.1. MSWEP V2.2

Multi-Source Weighted-Ensemble Precipitation (MSWEP) uses a combination of the gauge-, satellite-, and reanalysis-based data to provide a reliable precipitation estimates over the entire globe. With the consistent precipitation record from 1979 until the near present, it enables assessing the precipitation trend, drought, etc. In this study, the newest version of this product (V2.2) is used. The dataset cover the time period from 1979 to 2017 and is available at 0.1° spatial resolution and at three-hourly, daily, and monthly temporal resolutions [17]. The historic MSWEP data are freely available via www.gloh2o.org.
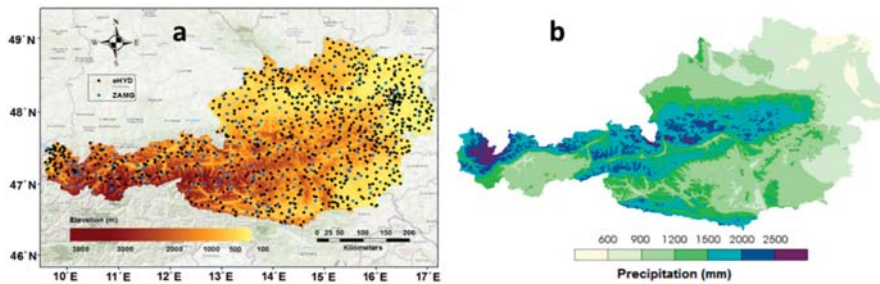
Figure 1. (**a**) Distribution of eHYD and ZAMG precipitation stations and (**b**) mean annual precipitation over Austria.

2.2.2. ERA5

The precipitation data obtained from ERA5 forecast product developed through the Copernicus Climate Change Service (C3S). This product has global spatial coverage and covering the period 1950 to the present. Precipitation data are available at an hourly resolution and consist of forecasts initialized twice daily from analyses at 06 and 18 UTC. The ERA5 high-resolution atmospheric data has a resolution of 0.28125 degrees (31 km) but, when downloading the data, there is a possibility to resampling the data to a higher spatial resolution. This will be done for continuous parameters by default through bilinear resampling methods [18]. Therefore, for consistency to other satellite data, we used the precipitation data with 0.1° spatial resolution. The data are freely accessible to users from Climate Data Store (CDS) website (https://cds.climate.copernicus.eu).

2.2.3. IMERG-FR (Final-Run) V05B

The Global Precipitation Measurement (GPM) mission's Precipitation Processing System (PPS) at NASA's Goddard Space Flight Center released IMERG-V05B data to the public in late November 2017. The dataset includes precipitation since March 2014. IMERG datasets are freely accessible to users from NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) website (https://disc.gsfc.nasa.gov). The IMERG-FR product is produced based on merging and interpolating all constellation microwave sensors and IR-based observations, and calibrated with monthly precipitation rain gauges from GPCC and create research-level products [19].

2.2.4. IMERG-RT (Real-Time) V05

The data are available from March 2014 to the present, thus providing IMERG-RT V03 data for the period 2014, to December 2015. The IMERG-RT late run, which is the near real-time product of the IMERG. It runs about ~10–14 h after the observation time and it is calibrated with the monthly climatological data (unlike the Final-Run, which uses monthly gauge data).

2.2.5. IMERG-FR (Final-Run) V06A

In the new version of IMERG, datasets from the Tropical Rainfall Measuring Mission (TRMM) have been used as a calibrator for GPM for the nearly time period of TRMM start its observation, to allow GPM processing spin up and a graceful transition from the TRMM era to the GPM era. This product was released on 26 March 2019 and the first available GPM-era products start with June 2014 until December 2016 and soon it will span 2000 until the present [20].

2.2.6. SM2RAIN-ASCAT

SM2RAIN obtained from ASCAT (Advanced SCATterometer) satellite soil moisture data [21]. The SM2RAIN-ASCAT rainfall dataset is provided over an irregular grid at 12.5 km on a global scale. The

product represents the daily cumulated rainfall. The SM2RAIN method was applied to the ASCAT soil moisture product for the period from 2007 to 2018.

3. Methodology

For our study period, the 782 eHYD gauges were used in this domain. The eHYD accumulates daily precipitation ending at 07:00 o'clock local time, which is different from the satellite daily precipitation accumulation convention. Therefore, the sub-daily precipitation estimates (e.g., half-hourly, three-hourly) were aggregated and derived a meaningful daily (at the local time of gauge measurement) and, consequently, monthly data. Then a comparison among the eHYD gauge-based data and the gridded precipitation products across Austria was conducted. It should be mentioned since the SM2RAIN data are available only in daily (and not sub-daily) scale, therefore, we could use only the original SM2RAIN daily data which estimate rainfall between the 00:00 and the 23:59 UTC of the indicated day.

In this research, to evaluate the capability of precipitation products to capture the precipitation patterns and for the aim of intercomparison with other gridded precipitation products, the data of a pixel of the gridded products are compared with that corresponding to the ground point observation (i.e., the station). Only the cells where there is at least one reporting station can be selected for computation. Since we have used a dense network, in numerous pixels, there are two or more gauges are located. In this case, an average of the two or more gauges is used as the basis for comparison. At the end, all stations fell within 601 pixels over the country.

The assessment and validation of precipitation products are carried out based on continuous and categorical statistical metrics. To quantitatively compare the performance of the gridded products against in situ observations at daily and monthly time-scales, the continuous statistical metrics including correlation coefficient (CC), bias, mean absolute error (MAE), and root mean square error (RMSE) are used. The CC is used to assess the agreement between SPEs and rain-gauge observations. The CC value vary from −1 to +1, where +1 indicates a perfect skill score and −1 indicates a perfect negative linear correlation. The bias is defined as the average difference between in situ observations and satellite/model precipitation estimates, and can be either positive or negative. A negative bias indicates underestimation by satellite precipitation while a positive bias indicates overestimation. The MAE is used to represent the average magnitude of the error and MAE = 0 indicates a perfect score. The RMSE is used to measure the average error magnitude and weighs the errors according to their squared value. This gives a greater weight to larger errors than the MAE. RMSE = 0 represents a perfect score. To examine the capability of the products in detection of precipitation, the two categorical statistical metrics, probability of detection (POD) and false alarm ratio (FAR), are used (see Appendix A). POD is an indicator of the SPE's ability to correctly detect precipitation events. Values vary from 0 to 1, with 1 as a perfect score. FAR denotes the fraction of cases in which the SPEs record precipitation when the rain gauges do not. Values vary from 0 to 1, with 0 as a perfect score. For extreme precipitation, we use R90th index to measure extreme wetter condition. R90 is precipitation in the 90th percentile of wet days in a year (i.e., after excluding precipitation less than 0.1 mm). We further break down and more deep analyses are conducted by classifying the stations' elevation (1000 m ≥ stations' elevation > 1000 m over the whole country).

The processing stages for error analysis of this study is shown in Figure 2.
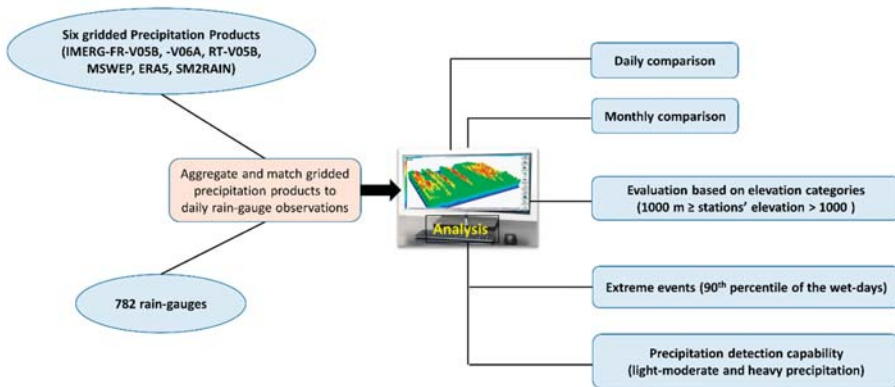
Figure 2. Processing stages for error analysis of this study.

4. Results

In the first step we evaluated the daily time-scale of IMERG-V05, -V05-RT, -V06A, MSWEP, ERA5, and SM2RAIN over all days from June 2014 to December 2015 against the in situ observations as reference. Considering that the IMERG-FR V06A started in June 2014 while the eHYD was discontinued by December 2015, the study period was confined to the 17 months between June 2014 and December 2015. It is worth to mention that further work is needed to evaluate the seasonal and inter-annual comparison of these products relying on larger sample data. Figure 3 shows the spatial distribution and average statistical indices (bias, CC, RMSE, and MAE) at daily precipitation time-scale for all products over Austria.

Precipitation types vary across the area. This region is typified by stratiform and convective precipitation, while the west and middle of the area (along with alpine mountains), in addition, is dominated by complex precipitation system due to the orography of the area. In Figures 3 and 4 the spatial distribution and the statistical summary of the metrics for the aforementioned products at a daily and monthly resolution over Austria are shown. According to Figure 3, precipitation shows a weaker correlation to ERA5 and SM2RAIN with mean CC of 0.53 and 0.57, respectively, in compare to MSWEP (CC = 0.86) at daily time-scale. Particularly, over the Alpine mountains both SM2RAIN and IMERG-V05-RT indicated low CC skills, while they showed a better CC over the east and middle of the country.

The CC metric is used to describe the agreement between gridded precipitation products and in situ observations. As can be seen in Figure 3, with respect to CC, MSWEP, significantly yields better than other products in the whole domain in the range of 0.8 to 1 in most pixels. However, ERA5 indicated very low CC over the south and northern part of the country and rather high CC in the area with low altitude. The general performances of the CC for all three versions of IMERG and SM2RAIN are relatively similar and relatively low over the western part, in comparison to MSWEP that several factors could contribute to this lower CC over such areas: a) the topography and climate of the west domain is partly complex, might rise a big challenge for SPE accuracy [12]; b) the GPCC stations that are used for the calibration of IMERG are in monthly time-scale, while in this study the examination of the products are on a daily time-scale, leads the quality of IMERG products being potentially degraded. In general, the MAE and RMSE are significantly higher in the high altitudes and low in low altitudes. This is due to more sensitivity of RMSE, and there were high number of local and heavy convective precipitation events over the high altitudes of Austria.

The general analysis of the results shows that IMERG-V05B, -RT, -V06A, and SM2RAIN and ERA5 have similar scores with respect to MAE and RMSE, although the ERA5 surpass the other products according to the bias skill scores. However, MSWEP indicates a better result according to the error indices with bias, RMSE and MAE of –15 mm, 2.86 mm and 1.08 mm, respectively. This means that in

Austria, MSWEP daily precipitation is very close to the in situ precipitation observations among the other recent and state-of-the-art precipitation products. These results are consistent with Beck et al. [22] which determine to underscore the importance of applying daily gauge corrections and accounting for gauge reporting times in compare to monthly gauge corrections. Meanwhile, IMERG and ERA5 indicate relatively higher variety of spatial bias and CC. However, one of the causes of the error of the gridded precipitation products might be their precipitation estimation for a whole pixel once there is a localized precipitation event, particularly in the west part of the area which characterized by complex systems, in some of the stations within. Thereby wrongly assigning the event to unaffected stations [23]. The tendency of reanalysis data to overestimate precipitation frequency might be the cause of ERA5 precipitation overestimation [18,24]. Therefore, after numerous occurrences, this process causes an average areal overestimation/underestimation.

Although all the products exhibited almost similar mean statistical skill scores in overall, regionally there were considerable differences. Compared to MSWEP and IMERG-FR, ERA5, SM2RAIN, and IMERG-V05-RT performed substantially worse over regions of complex terrain [22]. The results suggest that the topography and climate characteristics of the region should be considered when choosing between satellite and reanalysis datasets.
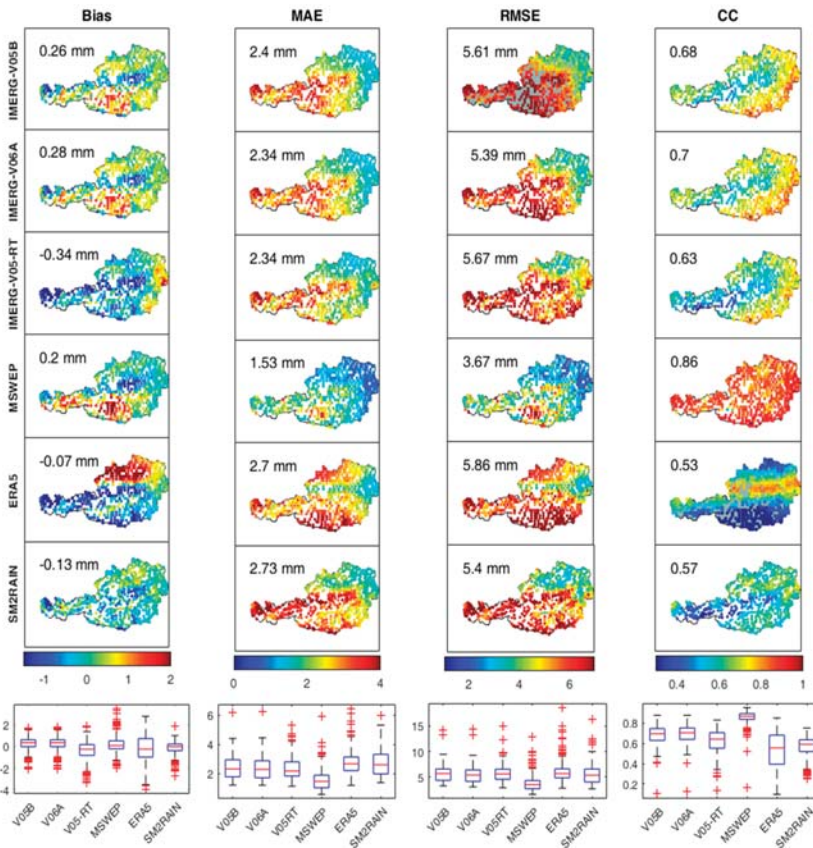
Figure 3. Statistical indices for bias, MAE, RMSE, and CC from left to right columns, respectively, at daily time scale for IMERG-V05, IMERG-V06A, IMERG-V05-RT, MSWEP, ERA5, and SM2RAIN. The center-line of each boxplot depicts the median value (50th percentile) and the box encompasses the 25th and 75th percentiles of the sample data, while the whiskers represent the extreme values, respectively.
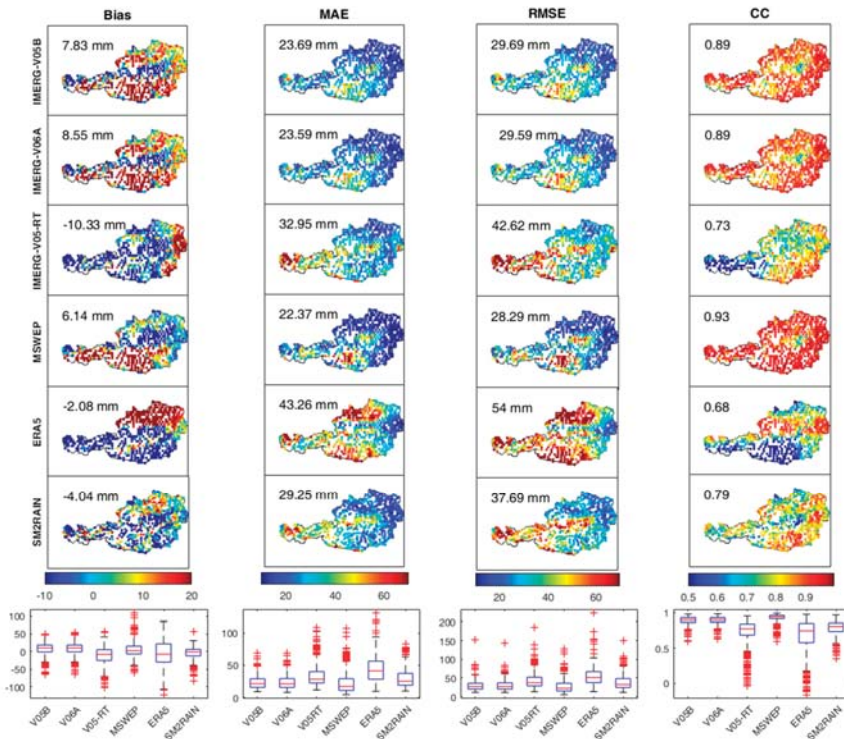
Figure 4. Statistical indices and for bias, CC, RMSE, and MAE from top to down rows, respectively, at monthly time scale for IMERG-V05B, IMERG-V06A, MSWEP, ERA5, and SM2RAIN. The center-line of each boxplot depicts the median value (50th percentile) and the box encompasses the 25th and 75th percentiles of the sample data, while the whiskers represent the extreme values, respectively.

The monthly statistical indices from all precipitation products versus in situ observations are shown in Figures 4 and 5. According to the results of monthly precipitation, although all the examined products indicated a rather close performance to in situ measurements, it is evident that MSWEP, followed by IMERG-V05B and -V06A monthly precipitation compared well to the corresponding in situ measurements. With respect to monthly scale, the CC of MSWEP, IMERG-V05B, -V06A, and SM2RAIN exhibited strong agreement with observations over the whole area. Although IMERG-V05-RT and ERA5 indicated rather good CC for the eastern part of the country, they showed weak performance for the western and southern parts of the region, respectively, which might be due to the effect of relief and complex systems in that area. MSWEP with the skill scores of 6.14 mm, 22.37 mm, 28.29 mm, and 0.93 and ERA5 with −2.08 mm, 43.26 mm, 54 mm, and 0.68 for bias, MAE, RMSE, and CC, respectively determined as the best and worst products. According to bias, the ERA5 strongly overestimated precipitation in the north part of the area and underestimated precipitation in south and west part of the country, with a mean areal bias value of −2.08 mm, that might be due to its native low-resolution and/or parameterization limitation during the precipitation generation processes [17,25]. The box-plots can confirm that most of the IMERG-V05B, -V06A, and MSWEP's pixels are in a smaller range, close to zero, in comparison to ERA5, with a wider bias range. This suggested that their gauge-correction methodology requires re-evaluation. Overall, MSWEP, followed by IMERG-V05B and -V06A, showed improvements in monthly precipitation in comparison with other products.
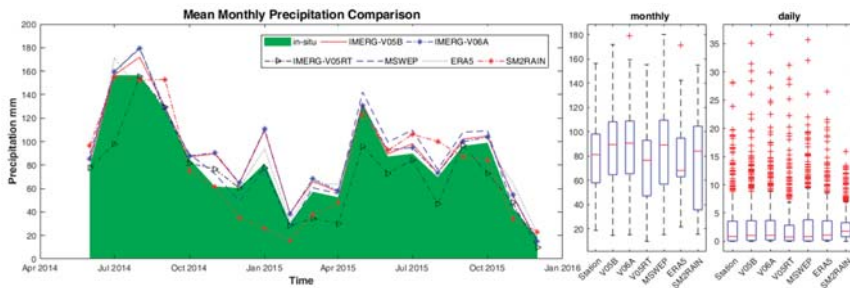
Figure 5. Time-series of mean monthly precipitation and box-plots of daily and monthly time-scales across Austria from IMERG, MSWEP, ERA5, and SM2RAIN products as compared to eHYD stations for the period of June 2014–December 2016. The center-line of each boxplot depicts the median value (50th percentile) and the box encompasses the 25th and 75th percentiles of the sample data, while the whiskers represent the extreme values, respectively.

Figure 5 shows the daily and monthly time-series comparison and boxplots of regional average precipitation from stations and other precipitation estimate products over Austria from June 2014 to December 2015. All precipitation products generally captured the spatio-temporal precipitation of daily and monthly time-scales, with the highest amount occurred in July and August 2014 and the lowest amount observed in February and December 2015. However, in monthly comparison it is evident that MSWEP estimates are very close to in situ observations and tend to slightly overestimate precipitation during August 2014 and July 2015, which might be due to the small scale of the precipitation systems that are dominant during these months, while IMERG-V05-RT seems to have systematic overestimation from March to November 2014.

According to Figure 5, MSWEP outperformed other products with slight overestimation over only the November 2014 and January 2015. The mean monthly data indicated that SM2RAIN underestimated precipitation during December 2014–April 2015 over Austria, which reflects a possible limitation of SM2RAIN-ASCAT data during the cold months. The SM2RAIN underestimation in winter can be related to snowfall that SM2RAIN is unable to estimate. The behavior of IMERG V05B and -V06A were almost similar with slight overestimation, while a greater overestimation of precipitation is observed mainly in the months with less precipitation intensity. IMERG-V05-RT shows systematic underestimation for February–October 2015. Additionally, from the median and the 25th and 75th percentiles of the box-plots, one can obtain that the precipitation estimated by MSWEP followed by IMERG-V05B and -V06A are more accurate than other products, although MSWEP whiskers extend to the most extreme data points. With respect to the box-plots of daily comparison, SM2RAIN indicated fewer extremes and outliers.
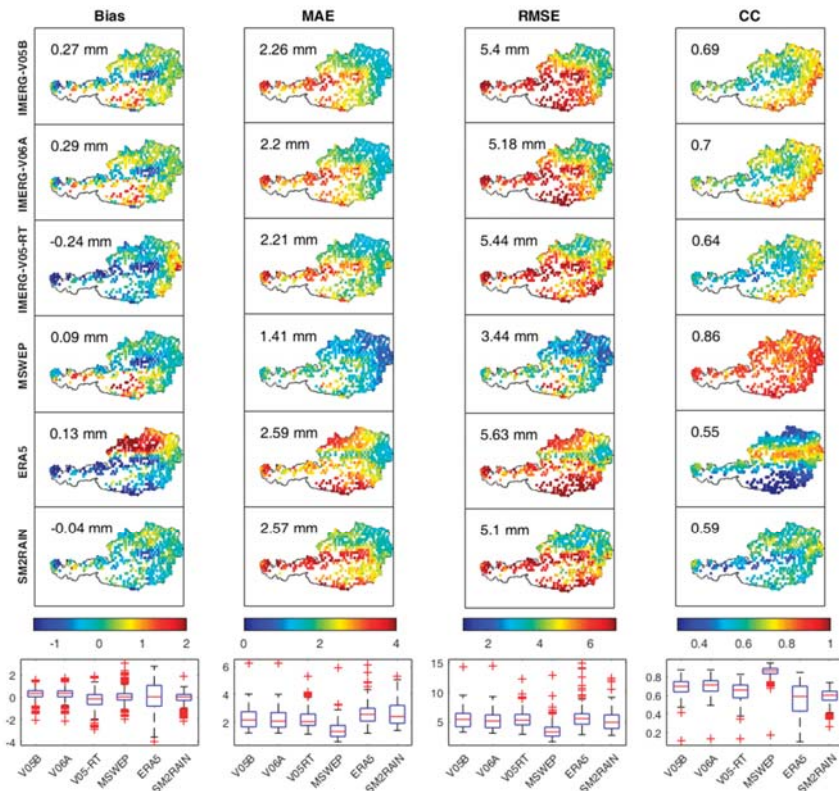
4.1. Elevation

Since Austria characterized by complex terrain and big difference in altitude over the country, annual mean precipitation range significantly varies with elevation and climate conditions. The microclimate can be created due to rapid changes in elevation which cause the obstruct the air mass movement or this rapid changes in elevation can cause the updraft of the air mass over the mountains to create orographic rainfall. Hence, for more deep analysis, the evaluation of the precipitation products was conducted by classifying the stations' elevation equal or less than 1000 m the stations located in greater than 1000 m altitudes (1000 m ≥ stations' Elevation > 1000 m) over the whole country in order to account for the effect of topography.

For the first category (1000 m ≥ Elevation), the performance of gridded precipitation products was evaluated by comparing daily data for 642 stations which fell into 502 pixels, while the second category contains 140 stations which fell into 125 pixels. Figures 6 and 7 show the spatial distribution of the

statistical indices for all products against in situ observations values for different elevation categories. MAE and RMSE evaluation metrics showed similar spatial patterns, while a sharp contrast from east to west of Austria for both elevation categories is observed, except for MSWEP, which indicated gradual variation. According to CC, MSWEP performed well, followed by IMERG-V05B and -V06A over the whole region, while ERA5, SM2RAIN, and IMERG-V05-RT showed weak CC, respectively, particularly over the alpine valleys.

With the increase of elevation, the mean RMSE, MAE, and bias increase and CC decreases, whereas the bias of IMERG-V05B and -V06A show a decreasing trend with increasing elevation/rainfall. The reason for this difference may be attributed to the cancellation of positive and negative biases, while logically due to high precipitation amounts the error should be higher. In other words, MAE and RMSE measure the absolute error magnitude and bias measure the relative error. The MAE, which evaluates the average magnitude error between precipitation products and in situ observations, were 2.26 mm, 2.2 mm, 2.21 mm, 1.44 mm, 2.59 mm, and 2.57 mm for IMERG-V05B, -V06A, -V05-RT, MSWEP, ERA5, and SM2RAIN, respectively, for the elevation category of less than 1000 m.

Figure 6. Spatial distributions and box plots of the statistical indices for the precipitation products and stations with the elevation equal or less than 1000 m. The center-line of each boxplot depicts the median value (50th percentile) and the box encompasses the 25th and 75th percentiles of the sample data, while the whiskers represent the extreme values, respectively.

Similarly, CC value of the aforementioned products was ≥ 0.5 in the majority of stations with an average value of 0.69, 0.70, 0.64, 0.86, 0.55, and 0.59, respectively, for the stations with less than 1000 m in altitude, while the average CC value of 0.64, 0.66, 0.55, and 0.85 obtained for the IMERG-V05B,

-V06A, -V05-RT, and MSWEP, respectively, for the stations located in the high altitudes. The ERA5 and SM2RAIN products failed to capture the observed daily precipitation with CC < 0.5 in most stations over the high altitudes and complex terrains. In general, one can say all products performed better in the low altitudes compared to the high altitudes.
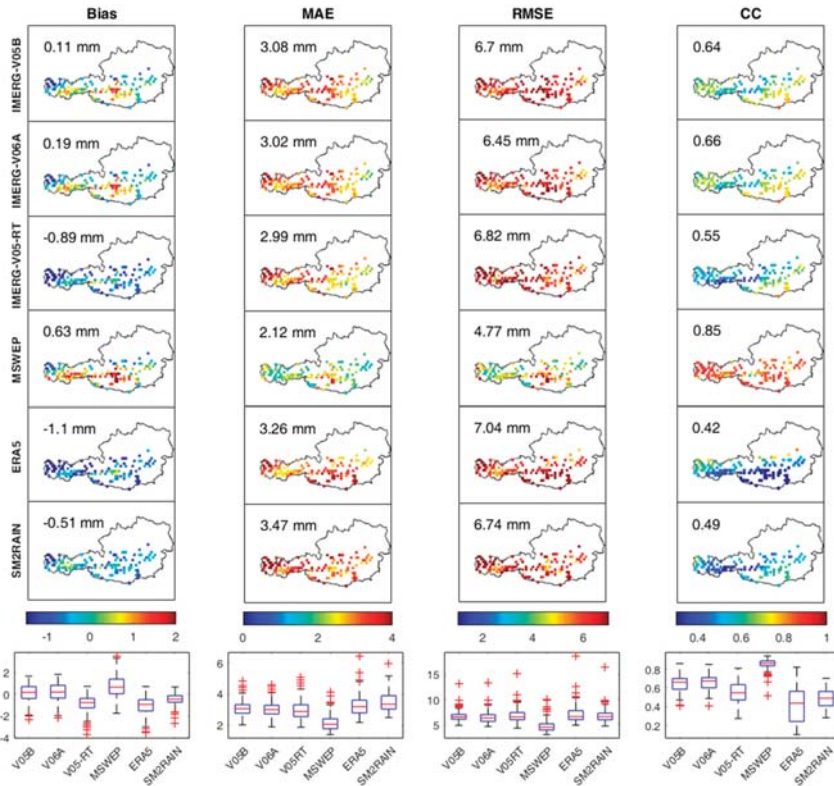
Figure 7. Spatial distributions and box plots of the statistical indices for the precipitation products and stations with the elevation greater than 1000 m. The center-line of each boxplot depicts the median value (50th percentile) and the box encompasses the 25th and 75th percentiles of the sample data, while the whiskers represent the extreme values, respectively.

It is notable to mention that inconsistent estimation of the precipitation products (except MSWEP) is possibly due to the rough terrains effect. The overall performance of the precipitation products is lower in the peripheries of the study area where most stations are situated in the mountainous area [26].

4.2. Extreme Events

In this part, R90th index using daily precipitation were examined to characterize the spatial distribution of daily precipitation and its extremes in order to cover the associated uncertainties. The 90% percentile level of wet days (P ≥ 0.1 mm) as the R90th threshold has been used. The resulting threshold for each station and precipitation products are shown in Figure 8. As can be seen, the lengths of extremes are double over the Alpine area in compare to low altitude regions (eastern part of the country). The stations' R90th showed the maximum values at high-elevation areas and in the west and northwest of the country. This region is considered the Alpine mountains with high mean annual precipitation amounts and has complex precipitation systems. The spatial distribution of the

R90th for MSWEP, IMERG-V05B, and –V06a were rather similar with higher number of days for the precipitation threshold above 90th percentile over the south part of the region, which showed the reliability of the estimations of this index. However, the spatial mean value of R90th for MSWEP was more close to the stations. In contrast, ERA5 underestimated extreme events over the big part of the south region, while showed higher number of extreme days over north Austria. Moreover, SM2RAIN, displayed underestimation of the R90th, almost over the whole country, except for some parts over the western region.
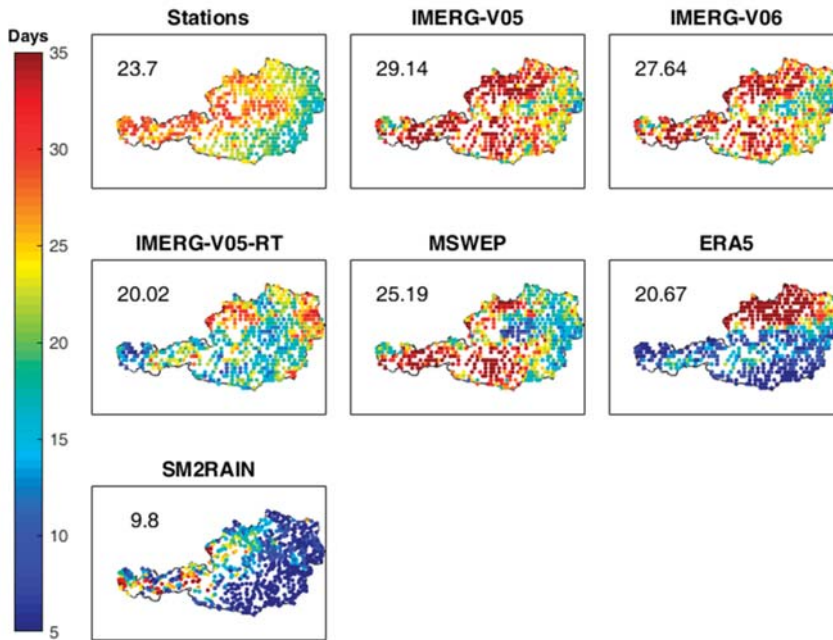
Figure 8. Distribution of daily R90th percentile of precipitation.

Figure 9 illustrates the spatial distributions of aggregated precipitation for stations (through bilinear interpolation), IMERG-V05B, -V06A, -V05-RT, MSWEP, ERA5, and SM2RAIN across Austria during the study period. The high precipitation areas extended from east to west along with alpine mountains. Although there are differences in magnitudes of precipitation among the products, in general, all products reasonably captured the precipitation distribution for most parts of the domain. The remarkable precipitation gradients are well-captured by MSWEP, possibly due to using daily in situ observation for bias correction in its algorithms when compared to IMERG, which uses monthly in situ observations for its bias correction. Another cause might be the native higher spatial resolutions of MSWEP (0.1° × 0.1°) than for example ERA5 product (~0.28° × ~0.28°). Nevertheless, ERA5 only poorly agrees with the gauge-based data at daily and monthly time scales, while patterns of accumulated precipitation agree well.

Moreover, IMERG indicates smoother precipitation trend from the west to the eastern part of the study area. The other reason for less comparable of IMERG products with MSWEP might be due to the limited temporal sampling of observations through active and passive microwave satellite sensors in comparison to MSWEP [27,28]. The station observation shows mean precipitation of 2.78 mm, whereas IMERG-V05B, -V06A, MSWEP, and ERA5 overestimate and IMERG-V05-RT underestimates over the whole domain with the mean precipitation values of 3.08 mm, 3.11 mm, 3.11 mm, 3.36 mm, and 2.41 mm, respectively.
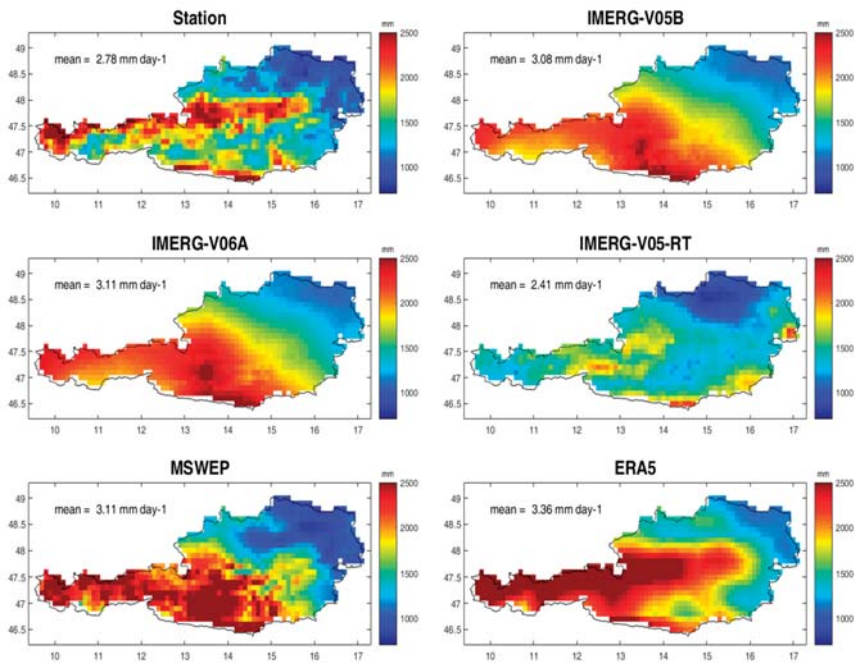
Figure 9. Spatial distribution of accumulated precipitation (mm) from June 2014 to December 2015 by stations, IMERG-V05B, IMERG-V06A, IMERG-V05-RT, MSWEP, and ERA5.

Figure 9 indicates that the total annual precipitation increases with elevation in the center of the country and extended to the west and south parts of the domain. In contrast, it reduces with elevation over the east parts. Stations located in the low altitudes of the eastern and northern parts of the basin receive less precipitation compared to the associated high altitudes.

4.3. Precipitation Detection Capability

The spatial distributions and box-plots of POD and FAR for the light-moderate precipitation range (0.1 mm ≤ P < 10 mm) and heavy precipitation (P > 10 mm) over Austria are shown in Figure 10. As can be seen, all products indicated acceptable skill scores in detecting light-moderate precipitation events. These results underscore the substantial advances in earth system modeling and SPE over the last decade. However, the POD of IMERG-05B has a mean areal value of 0.88, while that of IMERG-V06A has a mean value of 0.90, which shows an improvement of IMERG-V-06A over IMERG-V05B in detecting light-moderate precipitation events. Moreover, ERA5, MSWEP, and SM2RAIN indicated higher average POD with 0.94, 0.93, and 0.93, respectively, than all IMERG products. However, looking at the spatial distribution of POD indicating that the ECMWF's fourth-generation reanalysis (ERA5) and MSWEP have obvious advantages in detecting light-moderate precipitation events. In general, each value of POD of ERA5 and MSWEP is significantly better than other products, particularly IMERG products, at most of the stations, which might be due to the tendency of reanalysis data to overestimate light-moderate precipitation frequency [18,24]. In contrast, MSWEP is superior to other products to correctly detect precipitation and no-precipitation events.

The FAR for light-moderate precipitation range of all IMERG and MSWEP products presents almost similar spatial distribution pattern with better performance of MSWEP, particularly over the west part of the country. Figure 10 also shows that MSWEP has the lowest average FAR value (0.11) than other products over the area, while SM2RAIN and ERA5 reveal the highest FAR values (0.33 and 0.25).
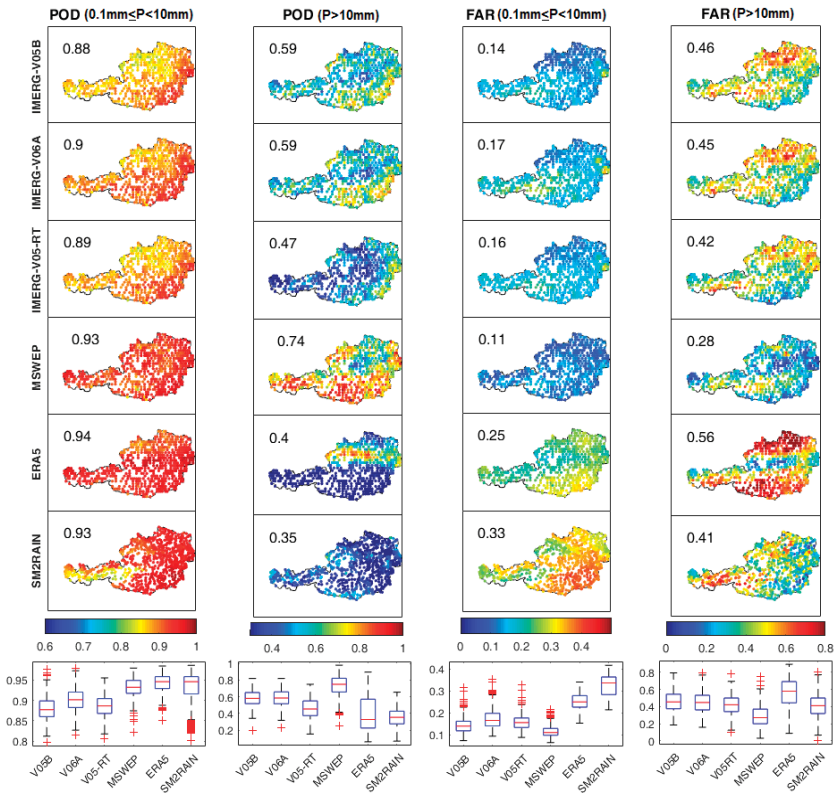
Figure 10. Spatial distributions and box-plots of POD and FAR at daily scale with respect to precipitation range of light-moderate (0.1 mm ≤ P < 10 mm) and heavy (P > 10 mm) events over Austria. The red-line in the middle of the box-plots represent the median value, the lines above and below the box represent the 25th and 75th percentile values, respectively, while the whiskers represent the extreme values.

As can be seen, the FAR values of IMERG products gradually rising from east to west, which indicates that it is more likely to appear false alarm in areas with higher precipitation. In addition, higher FAR might be due to the high amount of moisture in the atmosphere in this area that the satellites observed, although precipitation did not occur because of the evaporation of raindrops before reaching the ground [29].

For the heavy precipitation category (P > 10 mm), MSWEP and SM2RAIN products were found as the most and least powerful products to detect precipitation with the average value of 0.74 and 0.28 for MSWEP and 0.35 and 0.41 for SM2RAIN with respect to POD and FAR values over the area. Compared with MSWEP, IMERG-V05B, and -V06A, the ERA5 product has more complex spatial non-uniformity of POD and FAR. The MSWEP was found to dominate in the east and north, while the ERA5 dominated in the west in detecting the events. Despite this, a low POD and a high FAR of IMERG products for heavy precipitation mean that they are not able to properly detect precipitation in their exact precipitation categories (P ≥ 10 mm), but they might be able to detect the amount of precipitation somewhat lower than the specified intensities [3].

Lower POD for IMERG-V05-RT, particularly over the high altitudes in the west part of the country may be associated with missed precipitation over this region. The missed precipitation may be caused not only by snow cover on the ground at higher altitudes but also precipitation originates from small-scale and short-lived convective systems.

Notice that in the northern region of the study areas the value of FAR for precipitation above 10 mm/day from all IMERG products was rather high, which could be due not only to evaporation and not falling the small and tiny raindrops of the observed liquid water in the atmosphere profiles during the warm seasons, but also short-term precipitation events are highly variable in space and time and might not be detected by rain gauges, while being detected by other gridded products.

5. Discussion

In this research, assessment and comparison of the aforementioned precipitation products have provided insights into how different errors vary with precipitation intensities, elevation, and climate zones. With respect to CC, MSWEP significantly yields better than other products in the whole domain in the range in most pixels. However, ERA5, followed by SM2RAIN, indicated low CC over the southern and western parts of the country and rather high CC in the area with low altitude (Figure 3). The similarities of both IMERG-V05B and -V06A products are very consistent across the scores in daily and monthly time-scales over the whole country. The southern part of the domain is characterized by high positive biases up to ±2 mm/day for IMERG-V05B, -V06A, and MSWEP, whereas bias of the northern and eastern parts are much lower. This can be due to the higher precipitation intensities in these regions, ranging from 1400 mm to 2600 mm during the time period of this study. Moreover, varied orography and complex precipitation processes might be the other reason for this high statistical errors. However, MSWEP significantly performs other products, followed by IMERG-V06 and -V05B products. Moreover, the low bias, RMSE, and MAE and high CC along the eastern part of the domain (Lower Austria) shows an interesting feature. A considerable underestimation of precipitation along the complex train at the southern and western parts of the domain that characterized the complex precipitation is common to IMERG-V05-RT and ERA5 products.

The average monthly data showed that SM2RAIN underestimated precipitation during the cold months over Austria (Figure 5). This underestimation in winter can be related to snowfall and/or frozen soil, which SM2RAIN is unable to estimate. This finding is consistent with those shown by Paredes-Trejo et al. [30], who evaluated the performance of SM2RAIN over Brazil. Moreover, the ERA5 and SM2RAIN products failed to capture the observed daily precipitation with CC < 0.5 in most stations over the high altitudes (elevation > 1000 m) and complex terrains. In general, one can say all products performed better in the low altitudes (elevation < 1000 m) compared to the high altitudes. As the ASCAT soil moisture product has severe limitations over frozen soil, snow-cover, rainforest, and complex topographical regions, SM2RAIN-ASCAT, which is derived based on the ASCAT soil moisture product, also has difficulties to estimate precipitation over these regions [21,31–33]. In general, MSWEP properly captured the precipitation gradients, most likely due to using daily in situ observation for bias correction in its algorithms while IMERG uses monthly in situ observation for its calibration (Figure 9). Another cause might be the higher native spatial resolution of MSWEP (0.1° × 0.1°) than for example the ERA5 product (~0.28° × ~0.28°), or higher temporal resolution of MSWEP (3-hourly) in compare to for example SM2RAIN (daily). The algorithm of MSWEP optimally merges the gauge, satellite, and reanalysis precipitation estimates combining the advantages of the different data sources. Moreover, even though at daily and monthly time scales, ERA5 only poorly agrees with the gauge-based data, patterns of accumulated precipitation agree well. The adequate representation of spatial accumulation patterns may be due to (i) the high number of observations assimilated in ERA5, and (ii) assimilate precipitation data from ground-based radar observations (2009 onwards), although ERA5 fails to place the precipitation in the correct areas, when compared with rain gauges.

With respect to detecting light-moderate precipitation events, ERA5, MSWEP, and SM2RAIN indicated higher average POD compared to all IMERG products. For the heavy precipitation threshold (P > 10 mm) MSWEP indicated the most robust and SM2RAIN were found as the less powerful product to detect precipitation with respect to POD and FAR values over the area (Figure 10). The results indicated IMERG-V05-RT (for the entire country, except eastern and southeastern regions), ERA5 the (entire country, except a narrow band in the upper middle of the domain which is extended from

east to west), and SM2RAIN (the entire country) are unreliable at detecting precipitation at heavy precipitation category. Regarding the SM2RAIN precipitation products, this can be attributed to soil moisture retrieval errors, which highly affected the estimation of the precipitation quality derived from the SM2RAIN algorithm. SM2RAIN implemented a static correction procedure for climatological correction based on a cumulative density function (CDF) and the ERA5 reanalysis data [34]. Moreover, ERA5 showed some limitation, with emphasis on heavy precipitation, which can additionally affect the quality of SM2RAIN. In, overall, SM2RAIN-ASCAT, ERA5, and IMERG-V05-RT still face a significant challenge to estimate the amount of precipitation, while MSWEP-V2.2 and IMERG-V05B-FR, and V06A-FR revealed good performance to accurately estimate and detect precipitation over Austria. Thus, these products can offer a valuable alternative to in situ measurements for operational use in various applications.

6. Conclusions

To elucidate the strengths and weaknesses of recently released gridded precipitation datasets, we conducted a comprehensive evaluation of the performance of IMERG-FR-V05B, -V06A, IMERG-V05B-RT, ERA5, SM2RAIN-ASCAT, and MSWEP-V2.2 at daily and monthly time-scales for Austria using a dense network of gauges (882 stations) as a reference. The evaluation was carried out based on continuous and categorical statistical metrics for the period June 2014–December 2015. Apart from standard evaluations, we also assessed their performance with respect to the elevation of the stations and extreme events. In agreement with earlier studies, skills vary with respect to elevation, land surface characteristics and snow-rain phase of precipitation. In summary, our study shows that:

1. At the daily time-scale, MSWEP shows highest agreement with the gauge-based data followed by IMERG-V06A and -V05A. IMERG-V05-RT, ERA5, and SM2RAIN precipitation products show weaker correlations.

2. The skill scores of both IMERG-V05B and -V06A are very similar at daily and monthly time-scales and the newer version of IMERG (-V06A) did not indicate a strong improvement over IMERG-V05B over this area.

3. All precipitation products generally capture the spatio-temporal precipitation patterns at daily and monthly time-scales, with the highest precipitation amounts observed in July and August 2014 and the lowest amount observed in February and December 2015.

4. At the monthly time-scale, MSWEP precipitation estimates are very close to in situ observations while slightly overestimating precipitation during August 2014 and July 2015. This might be due to the small scale of the precipitation systems that are dominant during this months, while IMERG-V05B, -V06A, and ERA5 seem to have systematic overestimation over the entire months.

5. MSWEP outperforms other products. The mean monthly data indicate that SM2RAIN underestimates precipitation during the cold months which might be due to the snowfall that SM2RAIN is unable to estimate. The behavior of IMERG-V05B and -V06A was almost similar with slight overestimation in comparison with in situ observations, while a greater overestimation of precipitation is observed mainly in the months with less precipitation.

6. The overestimation or underestimation over the area with complex precipitation systems reveals that the newest generation of the satellite and reanalysis precipitation products examined in this study, even though they are improved, still have difficulties in estimating accurate precipitation amounts in capturing the effects of topography on precipitation.

7. For the heavy precipitation category (P > 10 mm), MSWEP and SM2RAIN products were found as the most and least powerful products with the average value of 0.74 and 0.28 for MSWEP and 0.35 and 0.41 for SM2RAIN with respect to POD and FAR values over the area.

8. The FAR for light-moderate precipitation (0.1 mm ≤ P < 10 mm) range of all IMERG and MSWEP products presents almost similar spatial distribution patterns with slightly better performance of MSWEP, particularly over the western part of the country.

9.   Regarding the heavy precipitation (P ≥ 10 mm) detection, MSWEP and IMERG products were found as the most and least powerful products, respectively, with the average value of 0.70 and 0.22 for MSWEP and 0.59 and 0.47 for IMERG-V05B for POD and FAR values over the area. Such robustness may be rooted in applying the daily gauge corrections for MSWEP.

10. Inadequate number of gauges, provided by the GPCC and using the monthly gauge correction might be a cause for IMERG products being less powerful compared to MSWEP, which uses daily gauge data.

11. The spatial distribution of the R90th for MSWEP, IMERG-V05B, and -V06a were rather similar, which showed the reliability of the estimations of this index. However, the spatial mean value of R90th for MSWEP was closer to the stations. In contrast, ERA5 underestimated extreme events over large part of the south of the country, while showing higher number of extremes over north Austria. Moreover, SM2RAIN underestimates the R90th, almost over the entire country, except for some parts over the western region.

12. According to the elevation categories, MAE and RMSE evaluation metrics showed almost similar skills for all products, while a sharp contrast between the east and west of Austria with respect to both elevation categories is observed, except for MSWEP, which indicated gradual variation. With respect to CC, MSWEP performed well, followed by IMERG-V05B and -V06A over the whole region, while ERA5, SM2RAIN, and IMERG-V05-RT showed weak CC, respectively, particularly over the alpine valleys.

13. Daily CC values of the all products were greater than 0.5 (CC ≥ 0.5) in the majority of stations with an average value of 0.69, 0.70, 0.64, 0.86, 0.55, and 0.59, for the stations with less than 1000 m in altitude, while the average CC value of 0.64, 0.66, 0.55, and 0.85 obtained for IMERG-V05B, -V06A, -V05-RT, and MSWEP, respectively, for the stations located at high altitudes (stations located above 1000 m asl). The ERA5 and SM2RAIN products failed to capture the observed daily precipitation with CC < 0.5 for most stations at high altitude and in complex terrain.

As expected, all products are able to reproduce the main characteristics of the precipitation in Austria. However, MSWEP performed significantly better than other products, followed by IMERG-V05B and -V06A. Except for MSWEP, the other products indicated difficulty capturing the effects of relief on precipitation over the complex terrain. This research is to our knowledge the first study to evaluate IMERG-V06A, ERA5, MSWEP-V02.2, and SM2RAIN-ASCAT over Austria. Since MSWEP-V2.2 is more robust, statistically, and has long-recorded data (from 1979–2017), we suggest using this product for further studies in climate applications over this region. Moreover, the IMERG data are also available from 2000 to the near present. The inclusion of elevation effects seems to be crucial for a realistic estimation of the spatial distribution of precipitation in mountains areas. We suggest the developers of the examined products highly consider the actual topography, steep terrain, and deep, narrow valleys, over the mountainous area to obtain more realistic precipitation amounts. Moreover, using daily precipitation gauge data, like MSWEP, instead of monthly precipitation data, like IMERG products, can significantly improve the accuracy of the precipitation products. It is worthy to mention that the SM2RAIN-ASCAT product is available only at a daily time step (00:00 UTC), while other products are available at sub-daily time-scales. Therefore, SM2RAIN-ASCAT was evaluated against rain gauges and other products with a few hours' difference, which may contribute to increased errors in this study. Moreover, in the case of gauge measurements, there might also be some uncertainties and systematic errors due to wind effects, wetting, evaporation, and splashing, which typically amounts to 5%–10% in summer, and even higher in snow conditions [35], which should be considered in further studies. Moreover, we suggest assessing the performance of the IMERG-V06B, which was not yet available during the analysis of this study, and the near-real-time product of MSWEP and use them in hydro-meteorological models. We expect that the regional analysis of the recently released gridded precipitation products revealed in this study can give users a broader perspective and understanding of the features associated with currently precipitation products. It is worth mentioning that the

performance ranking of the products may differ across the other regions depending on the amount of gauge data utilized and the quality control applied for each dataset as well.

Author Contributions: Conceptualization: E.S.; J.E. and W.D.; methodology: E.S.; analysis: E.S.; writing—original draft: E.S.; writing—review and editing: E.S.; J.E. and W.D.

Appendix A

The effectiveness of precipitation estimations was measured via the following metrics: the root mean square error (RMSE), bias, the mean absolute error (MAE), and the correlation coefficient (CC):

$$Bias = \frac{\sum_{i=1}^{n}(P_{S_i} - P_{O_i})}{N} \quad (mm), \tag{A1}$$

$$CC = \frac{\sum_{i=1}^{N}(P_{S_i} - \overline{P}_S)(P_{O_i} - \overline{P}_O)}{\sqrt{\sum_{i=1}^{N}(P_{S_i} - \overline{P}_S)^2}\sqrt{\sum_{i=1}^{N}(P_{O_i} - \overline{P}_O)^2}}, \tag{A2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(P_{S_i} - P_{O_i})^2} \ (mm), \tag{A3}$$

$$MAE = \frac{\sum_{i=1}^{N}|P_{S_i} - P_{O_i}|}{N} \ (mm), \tag{A4}$$

where P_{S_i} and P_{O_i} are the value of satellite/reanalysis precipitation estimates and the value of rain-gauge observations, respectively; i is the index of the station number and N the total number of stations; $\overline{P}_S$ and $\overline{P}_O$ are the average value of satellite precipitation estimates and rain-gauge observations for N stations over the study area.

Another assessment technique of satellite/reanalysis precipitation estimation is using a contingency table that reflects the frequency of "Yes" and "No" of the precipitation estimation products

Table A1. Contingency table.

	I/J	Observed		
		Yes	**No**	**Total**
Satellite	**Yes**	*Hit* (*a*)	*False alarm* (*b*)	*a* + *b*
	No	*Miss* (*c*)	*Correct negative* (*d*)	*c* + *d*
	Total	*a* + *c*	*b* + *d*	*n* = *a* + *b* + *c* + *d*

A dichotomous estimate says, "Yes, an event will happen", or "No, the event will not happen". By using this table for daily precipitation, a set of statistical indices are shown as follows:

Probability of detection (POD) responds to the question of what fraction of the observed "Yes" events were correctly estimated/forecasted. The perfect score is 1:

$$POD = \frac{hits}{hits + misses}, \tag{A5}$$

False alarm ratio (FAR) deals with the question of what fraction of the estimated/forecasted "Yes" events did not occur. The ideal score is 0:

$$FAR = \frac{false\ alarms}{hits + false\ alarms},$$ (A6)

References

1. Salimi, A.H.; Masoompour Samakosh, J.; Sharifi, E.; Hassanvand, M.R.; Noori, A.; von Rautenkranz, H. Optimized Artificial Neural Networks-Based Methods for Statistical Downscaling of Gridded Precipitation Data. *Water* **2019**, *11*, 1653. [CrossRef]
2. Sharifi, E.; Steinacker, R.; Saghafian, B. Assessment of GPM-IMERG and Other Precipitation Products against Gauge Data under Different Topographic and Climatic Conditions in Iran: Preliminary Results. *Remote Sens.* **2016**, *8*, 135. [CrossRef]
3. Sharifi, E.; Steinacker, R.; Saghafian, B. Multi time-scale evaluation of high-resolution satellite-based precipitation products over northeast of Austria. *Atmos. Res.* **2018**, *206*, 46–63. [CrossRef]
4. Berhan, G.; Hill, S.; Tadesse, T.; Atnafu, S. Using Satellite Images for Drought Monitoring: A Knowledge Discovery Approach. *J. Strateg. Innov. Sustain.* **2011**, *7*, 135–153.
5. Sharifi, E.; Saghafian, B.; Steinacker, R. Copula-based stochastic uncertainty analysis of satellite precipitation products. *J. Hydrol.* **2019**, *570*, 739–754. [CrossRef]
6. Beck, H.E.; van Dijk, A.I.J.M.; Levizzani, V.; Schellekens, J.; Miralles, D.G.; Martens, B.; de Roo, A. MSWEP: 3-hourly 0.25°; global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 589–615. [CrossRef]
7. Khodadoust Siuki, S.; Saghafian, B.; Moazami, S. Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products. *Int. J. Remote Sens.* **2017**, *38*, 558–571. [CrossRef]
8. Cattani, E.; Merino, A.; Levizzani, V. Evaluation of Monthly Satellite-Derived Precipitation Products over East Africa. *J. Hydrometeor* **2016**, *17*, 2555–2573. [CrossRef]
9. Dezfuli, A.K.; Ichoku, C.M.; Huffman, G.J.; Mohr, K.I.; Selker, J.S.; van de Giesen, N.; Hochreutener, R.; Annor, F.O. Validation of IMERG Precipitation in Africa. *J. Hydrometeor* **2017**, *18*, 2817–2825. [CrossRef]
10. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116. [CrossRef]
11. Hong, Y.; Tang, G.; Ma, Y.; Huang, Q.; Han, Z.; Zeng, Z.; Yang, Y.; Wang, C.; Guo, X. Remote Sensing Precipitation: Sensors, Retrievals, Validations, and Applications. In *Observation and Measurement*; Li, X., Vereecken, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 1–23.
12. Gottschalck, J.; Meng, J.; Rodell, M.; Houser, P. Analysis of Multiple Precipitation Products and Preliminary Assessment of Their Impact on Global Land Data Assimilation System Land Surface States. *J. Hydrometeor* **2005**, *6*, 573–598. [CrossRef]
13. Dinku, T.; Chidzambwa, S.; Ceccato, P.; Connor, S.J.; Ropelewski, C.F. Validation of high-resolution satellite rainfall products over complex terrain. *Int. J. Remote Sens.* **2008**, *29*, 4097–4110. [CrossRef]
14. Hofstra, N.; Haylock, M.; New, M.; Jones, P.; Frei, C. Comparison of six methods for the interpolation of daily, European climate data. *J. Geophys. Res.* **2008**, *113*. [CrossRef]
15. Thaler, S.; Brocca, L.; Ciabatta, L.; Eitzinger, J.; Hahn, S.; Wagner, W. Effects of Different Spatial Precipitation Input Data on Crop Model Outputs under a Central European Climate. *Atmosphere* **2018**, *9*, 290. [CrossRef]
16. Sharifi, E.; Saghafian, B.; Steinacker, R. Downscaling Satellite Precipitation Estimates with Multiple Linear Regression, Artificial Neural Networks, and Spline Interpolation Techniques. *J. Geophys. Res. Atmos.* **2019**, *124*, 789–805. [CrossRef]
17. Beck, H.E.; Wood, E.F.; Pan, M.; Fisher, C.K.; Miralles, D.G.; van Dijk, A.I.J.M.; McVicar, T.R.; Adler, R.F. MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Am. Meteorol. Soc.* **2018**. [CrossRef]
18. Skok, G.; Žagar, N.; Honzak, L.; Žabkar, 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]

19. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation, IMERG Tech Document. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf (accessed on 22 March 2017).

20. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Tan, J. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc_190313.pdf (accessed on 13 March 2019).

21. Brocca, L.; Filippucci, P.; Hahn, S.; Ciabatta, L.; Massari, C.; Camici, S.; Schüller, L.; Bojkov, B.; Wagner, W. SM2RAIN-ASCAT (2007–2018): Global daily satellite rainfall from ASCAT soil moisture. *Earth Syst. Sci. Data Discuss.* **2019**, 1–31. [CrossRef]

22. Beck, H.E.; Pan, M.; Roy, T.; Weedon, G.P.; Pappenberger, F.; van Dijk, A.I.J.M.; Huffman, G.J.; Adler, R.F.; Wood, E.F. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci. Discuss.* **2018**, 1–23. [CrossRef]

23. Quintana-Seguí, P.; Turco, M.; Herrera, S.; Miguez-Macho, G. Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-Interim. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 2187–2201. [CrossRef]

24. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.J.M.; Weedon, G.P.; Brocca, L.; Pappenberger, F.; Huffman, G.J.; Wood, E.F. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6201–6217. [CrossRef]

25. Herold, N.; Alexander, L.V.; Donat, M.G.; Contractor, S.; Becker, A. How much does it rain over land? *Geophys. Res. Lett.* **2016**, *43*, 341–348. [CrossRef]

26. Gebremicael, T.G.; Mohamed, Y.A.; van der Zaag, P.; Berhe, A.G.; Haile, G.G.; Hagos, E.Y.; Hagos, M.K. Comparison and validation of eight satellite rainfall products over the rugged topography of Tekeze-Atbara Basin at different spatial and temporal scales. *Hydrol. Earth Syst. Sci. Discuss.* **2017**, 1–31. [CrossRef]

27. Dorigo, W.; Wagner, W.; Albergel, C.; Albrecht, F.; Balsamo, G.; Brocca, L.; Chung, D.; Ertl, M.; Forkel, M.; Gruber, A.; et al. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.* **2017**, *203*, 185–215. [CrossRef]

28. Gebregiorgis, A.S.; Kirstetter, P.-E.; Hong, Y.E.; Gourley, J.J.; Huffman, G.J.; Petersen, W.A.; Xue, X.; Schwaller, M.R. To What Extent is the Day 1 GPM IMERG Satellite Precipitation Estimate Improved as Compared to TRMM TMPA-RT? *J. Geophys. Res. Atmos.* **2018**, *53*, 4434. [CrossRef]

29. Nashwan, M.S.; Shahid, S.; Wang, X. Assessment of Satellite-Based Precipitation Measurement Products over the Hot Desert Climate of Egypt. *Remote Sens.* **2019**, *11*, 555. [CrossRef]

30. Paredes-Trejo, F.; Barbosa, H.; dos Santos, C.A.C. Evaluation of the Performance of SM2RAIN-Derived Rainfall Products over Brazil. *Remote Sens.* **2019**, *11*, 1113. [CrossRef]

31. Prakash, S. Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *J. Hydrol.* **2019**, *571*, 50–59. [CrossRef]

32. Wagner, W.; Hahn, S.; Kidd, R.; Melzer, T.; Bartalis, Z.; Hasenauer, S.; Figa-Saldaña, J.; de Rosnay, P.; Jann, A.; Schneider, S.; et al. The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications. *Metz* **2013**, *22*, 5–33. [CrossRef]

33. Ciabatta, L.; Massari, C.; Brocca, L.; Gruber, A.; Reimer, C.; Hahn, S.; Paulik, C.; Dorigo, W.; Kidd, R.; Wagner, W. SM2RAIN-CCI: A new global long-term rainfall dataset derived from ESA CCI soil moisture. *Earth Syst. Sci. Data* **2018**, *10*, 267–280. [CrossRef]

34. Brocca, L.; Crow, W.T.; Ciabatta, L.; Massari, C.; de Rosnay, P.; Enenkel, M.; Hahn, S.; Amarnath, G.; Camici, S.; Tarpanelli, A.; et al. A Review of the Applications of ASCAT Soil Moisture Products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2285–2306. [CrossRef]

35. Haiden, T.; Kann, A.; Wittmann, C.; Pistotnik, G.; Bica, B.; Gruber, C. The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. *Weather Forecast.* **2011**, *26*, 166–183. [CrossRef]

Article

Raindrop Size Distributions and Rain Characteristics Observed by a PARSIVEL Disdrometer in Beijing, Northern China

Lei Ji [1], Haonan Chen [2,3,*], Lin Li [1,4], Baojun Chen [5], Xian Xiao [1], Min Chen [1] and Guifu Zhang [6]

[1] Institute of Urban Meteorology (IUM), Beijing 100089, China; lji@ium.cn (L.J.); lilin@bj.cma.gov.cn (L.L.); xxiao@ium.cn (X.X.); mchen@ium.cn (M.C.)

[2] Cooperative Institute for Research in the Atmosphere, Fort Collins, CO 80523, USA

[3] NOAA Earth System Research Laboratory, Boulder, CO 80305, USA

[4] Beijing Meteorological Observation Center, Beijing 100091, China

[5] Key Laboratory of Mesoscale Severe Weather (MOE), and School of Atmospheric Sciences, Nanjing University, Nanjing 210023, China; bjchen@nju.edu.cn

[6] School of Meteorology and Advanced Radar Research Center, University of Oklahoma, Norman, OK 73019, USA; guzhang1@ou.edu

* Correspondence: haonan.chen@colostate.edu; Tel.: +1-3034974616

Abstract: Fourteen-month precipitation measurements from a second-generation PARSIVEL disdrometer deployed in Beijing, northern China, were analyzed to investigate the microphysical structure of raindrop size distribution and its implications on polarimetric radar applications. Rainfall types are classified and analyzed in the domain of median volume diameter D_0 and the normalized intercept parameter N_w. The separation line between convective and stratiform rain is almost equivalent to rain rate at 8.6 mm h^{-1} and radar reflectivity at 36.8 dBZ. Convective rain in Beijing shows distinct seasonal variations in $\log_{10} N_w$–D_0 domain. X-band dual-polarization variables are simulated using the *T*-matrix method to derive radar-based quantitative precipitation estimation (QPE) estimators, and rainfall products at hourly scale are evaluated for four radar QPE estimators using collocated but independent rain gauge observations. This study also combines the advantages of individual estimators based on the thresholds on polarimetric variables. Results show that the blended QPE estimator has better performance than others. The rainfall microphysical analysis presented in this study is expected to facilitate the development of a high-resolution X-band radar network for urban QPE applications.

Keywords: Northern China; raindrop size distribution (DSD); microphysical processes; quantitative precipitation estimation (QPE)

1. Introduction

Characteristics of raindrop size distribution (DSD) are of great importance in various disciplinary research. They are the physical basis in the formation of clouds and precipitation [1]. Understanding the DSD is critical for the microphysical parameterizations in numerical weather prediction models [2–4], and quantitative precipitation estimation (QPE) using remote sensing technologies, such as radar and satellite [5,6]. The DSDs can also be utilized to estimate the kinetic energy of rain [7], which is a key factor in assessing the degree of soil erosion [8]. To this end, numerous studies have been conducted around the world to characterize the DSD in different climate regions and rainfall types, using a variety of in situ and remote sensing instruments [9–16].

The DSD can be affected by many factors [17], including microphysical processes, such as condensation, evaporation, collision–coalescence and breakup [18], updrafts and downdrafts [19], horizontal winds [20], orographic effects [21], and aerosol effects [22].

The climatological characteristics of precipitation in Beijing, China, have been examined using rainfall data collected at automatic weather stations [23,24] and radar reflectivity mosaics [25,26]. However, the microphysical structure of surface precipitation in Beijing is rarely reported, due to the lack of long-term ground-based DSD measurements. Using a first-generation laser-optical particle size and velocity (PARSIVEL) disdrometer manufactured by OTT Hydromet, Germany [27], Tang et al. [28] compared the characteristics of measured and fitted DSDs, as well as the retrieved dual-polarization radar variables for stratiform and convective precipitation in Beijing. However, the DSD samples used by Tang et al. [28] were only collected from July to October 2008, which did not include precipitation occurred in June that makes a significant contribution to the total annual rainfall in Beijing [24,29]. In addition, those DSD data were collected mainly under the conditions of improved air quality and lower aerosol concentration associated with strict emission-reduction during the Beijing Olympic and Paralympic Games [30], which may not be sufficient to represent normal air quality conditions in Beijing [31], since the concentrations and components of aerosols could potentially affect the DSD properties [22,32]. A second-generation PARSIVEL disdrometer (hereafter referred to as PARSIVEL^2) was used to study the snowfall properties over the mountains in northwestern Beijing [33]. Unfortunately, no long-term rainfall observations were reported using this instrument.

From 2017, a PARSIVEL^2 disdrometer was deployed at a national weather station in Beijing (116.47°E, 39.8°N; 31.3 m a.s.l.) to perform continuous microphysical measurements of rainfall on the ground, which provides an opportunity to investigate the characteristics of local DSD comprehensively. In addition, the DSD data can provide a means for improving the accuracy of remote sensing retrievals, such as polarimetric radar quantitative precipitation estimation (QPE) [34,35] and enhance the operational weather forecast model in Beijing (i.e., the Rapid-refresh Multi-scale Analysis and Prediction System–RMAPS [36]). This study aims to conduct a detailed investigation of DSD characteristics in Beijing using this disdrometer data. This paper is organized as follows. Section 2 describes the data and analysis methods, including the data quality control procedure and DSD parameters to be included in this study. Based on the quality-controlled disdrometer dataset, Section 3 describes the microphysical properties of DSDs in $\log_{10} N_w$–D_0 domain, as well as the comparison with other climate regions. Classification of different rain types is also detailed in Section 3. Section 4 derives the radar-based QPE estimators and quantifies the associated errors of various estimators using collocated gauge measurements. Major conclusions are summarized in Section 5.

2. Dataset, Quality Control, and DSD Parameters

2.1. Observations

The PARSIVEL^2 disdrometer is located at Beijing station in the North China Plain surrounded by the Yan mountains to the west and north, and the gulf-like Bohai Sea to the southeast (Figure 1). The mean annual precipitation was 575 mm during the most recent decade (2009–2018). In this study, 14 months continuous DSD measurements in 2017–2018, ranging from 1 April to 31 October in each year, were used, which made up 96.7% of the total rainfall (1085.8 mm out of 1122.7 mm) during this period. In order to focus on rainfall analysis, winter precipitation (mainly snowfall from November to March the next year), as well as the solid precipitation (such as hail), was removed according to the ground weather reports.
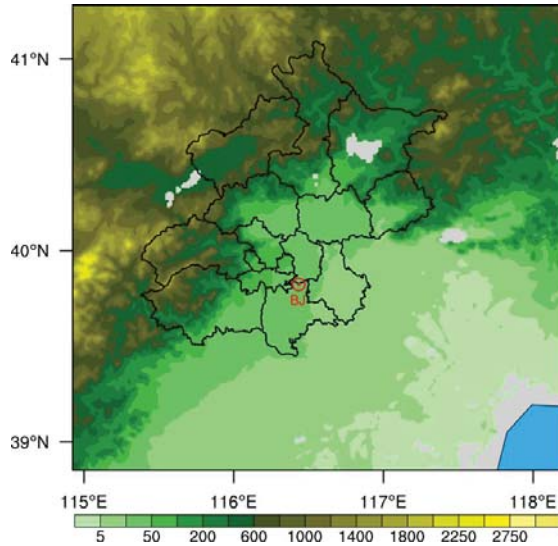
Figure 1. Topographic (m) information around the PARSIVEL[2] disdrometer site at Beijing station (BJ, the red circle). The districts of Beijing are highlighted in black curves.

2.2. Quality Control (QC)

Particle diameter and fall speed, each divided into 32 nonuniform classes, were measured by the PARSIVEL[2] disdrometer with a 1-min sampling interval. The mean values of particle diameter (0.062–24.5 mm) and fall speed (0.05–20.8 m s^{-1}) are described by the manual [37]. The first two size bins are not included in the analysis, because of the low signal-to-noise ratios. As a result, the smallest detectable mean diameter is 0.312 mm. The effective sampling area of PARSIVEL[2] droplet size measurements is affected by the so-called border effects, and the method of Jaffrain and Berne [38] is utilized to account for these effects. In particular, defining D_i (mm) as the central volume-equivalent diameter for the *i*th size bin, the effective sampling area can be calculated as 180 mm × (30 mm − 0.5D_i).

The empirical terminal velocity–diameter (*V*–*D*) relationship of Gunn and Kinzer [39] with air-density correction factor $(\rho_0/\rho_a)^{0.4}$ [40,41] was used to assess raindrop observations and is repeated as follows:

$$V_t(D_i) = [9.65 - 10.3\exp(-0.6D_i)]\left(\frac{\rho_0}{\rho_a}\right)^{0.4}, \tag{1}$$

where $V_t(D_i)$ is the mean particle terminal velocity for the *i*th size bin; ρ_a and ρ_0 (1.20 kg m^{-3}) are the air density at the observation altitude and at sea level, respectively. Following the method described in Atlas et al. [40] and Foote and Toit [41], the mean value (1.008) of the correction factor was selected for simplicity.

Some droplet observations may deviate from the *V*–*D* relationship shown in Equation (1). A commonly used method to eliminate those abnormal particles is to set a threshold regarding Equation (1). A value of ±60% was selected as the threshold [20] in this study, which means droplets with velocities of $V_{obs}(D_i)$ were discarded when they met the condition $\left|V_{obs}(D_i) - V_t(D_i)\right| > 0.6V_t(D_i)$. In addition, the 1-min DSD spectrum with a total number of raindrops $C_T < 10$ or a rain rate lower than 0.01 mm h^{-1} was considered to have no rain. Rain drops larger than 8 mm in diameter were also removed. Then, continuous spectra with rain-free periods of no longer than 1 h were defined as a rain event, and rain events lasting less than 5 min were eliminated to reduce the statistical errors. The dataset after quality control is further described in Section 3.1.

2.3. Integral Rainfall Parameters

Based on the DSD data, the number concentration of raindrops per unit volume per unit diameter interval for the ith size bin, $N(D_i)$ (m^{-3} mm^{-1}), can be calculated using Equation (2):

$$N(D_i) = \sum_{j=1}^{32} \frac{n_{ij}}{A_i \cdot \Delta t \cdot V_j \cdot \Delta D_i},$$ (2)

where n_{ij} is the number of raindrops at the ith size bin and the jth velocity class; A_i (m^2) and ΔD_i (mm) are the effective sampling area and width of the diameter interval at size D_i; V_j (m s^{-1}) is the fall speed for the jth velocity class; and Δt is the sampling time interval, which was set to 60 s in this study.

To further understand the characteristics of rainfall, the integral parameters of total number concentration N_T (m^{-3}), rainwater content W (g m^{-3}), rain rate R (mm h^{-1}), median volume diameter D_0 (mm), mass-weighted mean diameter D_m (mm), normalized intercept parameter N_w (m^{-3} mm^{-1}), and mass spectrum standard deviation σ_m (mm), were also calculated as follows:

$$N_T = \sum_{i=1}^{32} \sum_{j=1}^{32} \frac{n_{ij}}{A_i \cdot \Delta t \cdot V_j},$$ (3)

$$W = \frac{\pi}{6} \times 10^{-3} \cdot \rho_w \cdot \sum_{i=1}^{32} \sum_{j=1}^{32} D_i^3 \frac{n_{ij}}{A_i \cdot \Delta t \cdot V_j},$$ (4)

$$R = 6\pi \times 10^{-4} \cdot \sum_{i=1}^{32} \sum_{j=1}^{32} D_i^3 \frac{n_{ij}}{A_i \cdot \Delta t},$$ (5)

$$\frac{1}{2} W = \frac{\pi}{6} \rho_w \cdot \int_0^{D_0} D^3 N(D) dD,$$ (6)

$$N_w = \frac{3.67^4}{\pi \rho_w} \left(\frac{10^3 W}{D_0^4} \right),$$ (7)

$$D_m = \frac{\sum_{i=1}^{32} N(D_i) \cdot D_i^4 \cdot \Delta D_i}{\sum_{i=1}^{32} N(D_i) \cdot D_i^3 \cdot \Delta D_i},$$ (8)

$$\sigma_m = \frac{\sum_{i=3}^{32} (D_i - D_m)^2 N(D_i) \cdot D_i^3 \cdot \Delta D_i}{\sum_{i=1}^{32} N(D_i) \cdot D_i^3 \cdot \Delta D_i},$$ (9)

where ρ_w is the water density (1.0 g cm^{-3}).

Considering the emerging development of X-band dual-polarization weather radar for urban hydrometeorological applications [42,43], a set of dual-polarization radar variables, including radar reflectivity in the horizontal (vertical) polarization Z_h (Z_v) (mm^6 m^{-3}), differential reflectivity Z_{DR} (dB) and specific differential phase K_{DP} (° km^{-1}), are derived from DSDs using the T-matrix scattering technique [44]:

$$Z_{h,v} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \sum_{i=1}^{32} \left| f_{hh,vv}(D_i) \right|^2 N(D_i) \Delta D_i,$$ (10)

$$Z_{DR} = 10 \log_{10} \left(\frac{Z_h}{Z_v} \right),$$ (11)

$$K_{DP} = \frac{180\lambda}{\pi} \sum_{i=1}^{32} Re[f_{hh}(0, D_i) - f_{vv}(0, D_i)] N(D_i) \Delta D_i,$$ (12)

where $f_{hh,vv}(D_i)$ is the backscattering amplitude of a droplet with horizontal and vertical polarization; $f_{hh}(0, D_i)$ and $f_{vv}(0, D_i)$ are the standard forward scattering amplitudes, which is related to the depolarization factor and relative permittivity of water dielectric [45]; K_w is the dielectric factor of water (0.9639); and λ (mm) is the radar wavelength (3 cm). Note that Z_h (Z_v) in the unit of mm^6 m^{-3} is replaced by Z_H (Z_V) in the unit of dBZ wherever required in this paper, and $Z_{H,V} = 10 \times \log_{10} Z_{h,v}$.

3. Results

3.1. Dataset after QC

In total, 25,499 (934) 1-minute raindrop spectra passed (failed) the QC. The validated spectra account for a total rainfall of 1013.78 mm. According to the histogram in Figure 2, DSD samples failed to pass the QC mainly appear when rain rates (R_{stn}) measured by collocated rain gauges at 1-min-interval were lower than 15 mm h^{-1}. Falling beyond the threshold of the empirical *V–D* relationship is the major factor leading to droplet removal from the dataset, and accounts for 3.2% of total rainfall. It was also noted that most of the removed DSD samples were characterized by abnormally rain rates (R) compared with R_{stn}, most of which occurred when $R_{stn} < 10$ mm h^{-1} or $R_{stn} > 100$ mm h^{-1} (red points in the scatter plot of Figure 2). The Pearson correlation coefficient (PCC) between the pairs of (R, R_{stn}) was higher after QC (0.96 vs. 0.91). The linear fitting curve based on the dataset with $R_{stn} > 0$ mm h^{-1} after QC (blue line; denoted "QC + $R_{stn} > 0$") is close to the diagonal line.
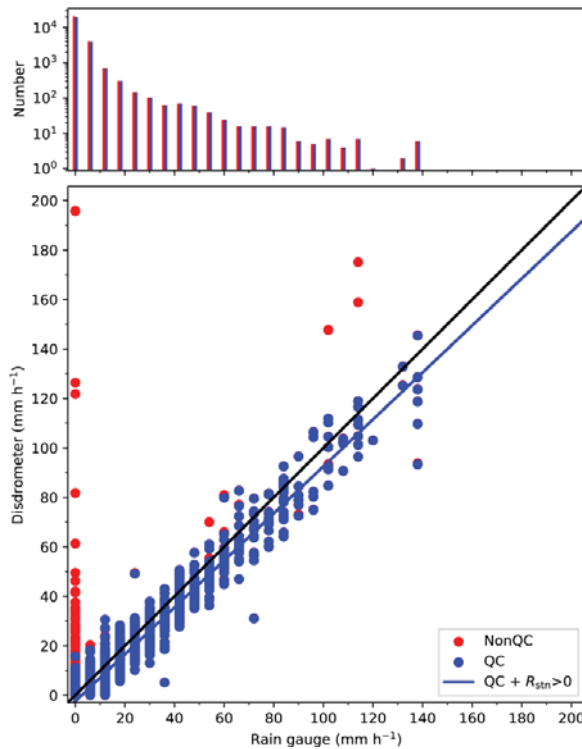
Figure 2. Histogram (top) of the number of 1-min raindrop spectra coinciding with rain gauge measurements (R_{stn}); and scatterplot (bottom) of rain rate calculated by PARSIVEL^2 disdrometer measurements vs R_{stn} observations from rain gauge at BJ during the experiment period. The solid black line in the scatterplot is the 1:1 line. Data before (NonQC) and after (QC) quality control are indicated by red and blue dots, respectively.

As shown in Figure 3, the distribution of raindrops is almost entirely within the threshold of ±60% based on Equation (1). The filtered particles are mainly below 3 mm in diameter. They generally have low fall speeds but with relatively large size, likely due to the influences of strong winds or splashes from instrument surface during heavy rainfall [20]. The accumulated disdrometer data after QC are almost symmetric along the empirical *V*–*D* relationship of Atlas et al. [40] and the highest number concentrations of raindrops are nearly superimposed.
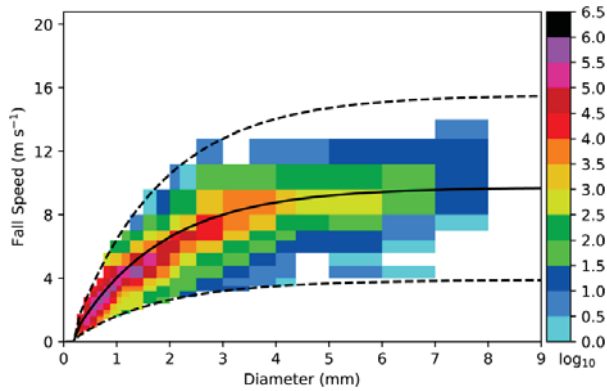
Figure 3. Scattergram of raindrop size distribution (DSD) at different diameter size and fall velocity classes after QC for the entire experiment period. The solid curve indicates the empirical *V*–*D* relationship described by Atlas et al. [40] which considers the air density effect; dashed curves indicate the ±60% ranges of the empirical *V*–*D* relationship.

A summary of rainfall observations after QC during the experiment period is listed in Table 1. The precipitation mainly occurred from June to August, which contributed up to 81.5% of the total rainfall amount. The mean and maximum rain rates, $\langle R \rangle$ and R_{max}, were much higher during these three months than other months. The number of DSD samples, N_{mins}, collected between June−August and in October, was much higher, contributing 78.3% of total samples. Although N_{mins} in October was higher than June, $\langle R \rangle$, R_{max}, and the rainfall amount were much lower in October, especially R_{max} (12.17 mm h^{-1} vs. 84.92 mm h^{-1}). The most (least) contribution of rainfall amount, as well as R_{max}, came from July (September), while the least $\langle R \rangle$ and N_{mins} came from April and September, respectively. Compared with 2017, the precipitation intensity in 2018 was heavier with higher $\langle R \rangle$ and R_{max} but lower N_{mins} and total rainfall amount. All these imply that the selected rainfall events consist of a wide variety of rainfall types.

Table 1. Summary of rainfall during the experiment period.

Type	April	May	June	July	August	September	October	2017	2018
N_{mins}	2599	1910	4374	5373	5396	1036	4811	14319	11180
$\langle R \rangle$ (mm h^{-1})	0.85	1.33	2.14	4.00	3.47	1.64	1.00	2.23	2.58
R_{max} (mm h^{-1})	26.46	45.72	84.92	145.43	123.61	10.02	12.17	118.92	145.43
Amount (mm)	36.63	42.48	155.88	358.28	312.44	28.26	79.80	532.78	481.00

Note: N_{mins} is the number of 1-min DSD samples. $\langle R \rangle$ and R_{max} are the mean and max rain rate, respectively.

3.2. Statistical Properties of N_w–D_0

N_w and D_0 are two main parameters defining the DSD [46,47], which also play an important role in retrieving precipitation microphysics on a global scale as part of the GPM mission [48,49]. In fact, major microphysical processes that dominate the DSD properties can partially be recognized in the $\log_{10} N_w$–D_0 domain [46]. The distribution of $\log_{10} N_w$ vs D_0 is also an indicator to separate convective

and stratiform rain types (C−S). In this study, the separation scheme described in Bringi et al. [50] (hereafter referred to as BR09) is adopted, as shown in Equation (13). Briefly, N_w–D_0 pairs above (below) Equation (13) are recognized as convective (stratiform) rain,

$$\log_{10} N_w^{\text{BR09}} = -1.6D_0 + 6.3. \tag{13}$$

By using C_BR09 and S_BR09 to, respectively, denote the convective and stratiform rain, classified by Equation (13), Table 2 summaries a series of DSD parameters for different rainfall types. There are 1488 (24011) minutes of DSDs classified as convective (stratiform) rain, which account for 5.8% (94.2%) of the entire dataset of occurance and correspond to 54.8% (45.2%) of total rainfall amount. Generally, the means of all DSD parameters for C_BR09 are higher than those for S_BR09.

Table 2. Properties of DSDs for different rain-type classification schemes.

Type	C_BR09	C_BR03	C_TE01	S_BR09	S_BR03	S_TE01
Spectra (min/%)	1488/5.8	1858/7.3	2134/8.4	24011/94.2	22094/86.6	23365/91.6
Amount (mm/%)	555.22/54.8	605.23/59.7	596.33/58.8	458.55/45.2	347.43/34.3	417.45/41.2
$\langle R \rangle$ (mm h^{-1})	22.39	19.54	16.77	1.15	0.94	1.07
1%/99% (mm h^{-1})	1.03/104.55	5.10/102.63	0.16/100.96	0.02/7.23	0.01/6.13	0.01/6.79
$\langle W \rangle$ (g m^{-3})	1.08	0.97	0.83	0.08	0.07	0.08
$\langle N_T \rangle$ (m^{-3})	1179.96	1132.05	1017.30	318.09	299.14	309.12
$\langle Z_H \rangle$ (dBZ)	43.24	41.40	38.58	19.62	18.82	19.39
$\langle Z_{DR} \rangle$ (dB)	1.75	1.48	1.34	0.38	0.36	0.38
$\langle K_{DP} \rangle$ (° km^{-1})	1.71	1.43	1.23	0.04	0.03	0.04
$\langle \log_{10} N_w \rangle$	3.61	3.77	3.72	3.57	3.56	3.56
$\langle D_0 \rangle$ (mm)	2.03	1.82	1.72	1.01	0.99	1.01
$\langle D_m \rangle$ (mm)	2.05	1.86	1.76	1.03	1.01	1.03
$\langle \sigma_m \rangle$ (mm)	0.78	0.70	0.66	0.32	0.31	0.32

Note: Rain types and classification schemes are listed in the first row. 'C'/'S' indicates convective/stratiform rain, whereas 'BR09', 'BR03' and 'TE01' represent the classification schemes developed by Bringi et al. [50], Bringi et al. [51], and Testud et al. [52], respectively. For example, C_BR09 and S_BR09 correspond to convective and stratiform rain classified by BR09 scheme. The number of spectra (occurrence), as well as their proportion of the entire dataset are given before and after the '/' in row 2. Row 3 is same as row 2, but for the rainfall amount. The 1th and 99th quantiles of rain rate for each dataset are listed before and after the '/' in row 5. Angle bracket stands for the sample mean.

Figure 4 shows the scatterplot of $\log_{10} N_w$ versus D_0 for convective (C_All, orange) and stratiform (S_All, lime) rain types, as well as the corresponding relative occurance frequency. The mean (MEAN), standard deviation (STD) and skewness (SKEW) are also indicated in Figure 4. Here, C_All (S_All) dataset equals to the dataset of C_BR09 (S_BR09) denoted in Table 2. Equation (13) are superimposed in the scatterplot panel (dashed line). Meanwhile, another C−S separation line suggested by Thompson et al. [53] (hereafter referred to as TH15) for oceanic, tropical rain regions is also superimposed (dot-dashed line) for reference. Equation (14) shows the formula of TH15,

$$\log_{10} N_w^{\text{TH15}} = 3.85. \tag{14}$$

Stratiform samples (S_All) are concentrated near the MEAN values of $D_0 = 1.01$ mm and $\log_{10} N_w = 3.57$, whereas convective samples (C_All) are sparsely distributed above the BR09 line. It results in larger STD of D_0 and $\log_{10} N_w$ for convective than stratiform rain. The D_0 histograms for both rain types are positively skewed, whereas the $\log_{10} N_w$ histograms for convective rain exhibit a negative skewness of −0.93. Compared with stratiform rain, the D_0 and $\log_{10} N_w$ histograms for convective rain tend to shift toward larger values, which are in agreement with previous studies for other climate regimes [10,11,51]. Similar variation tendencies of D_0 and $\log_{10} N_w$ histograms between "Total" dataset (blue) and stratiform rain can be found, which are due to the dominant role of stratiform rain.
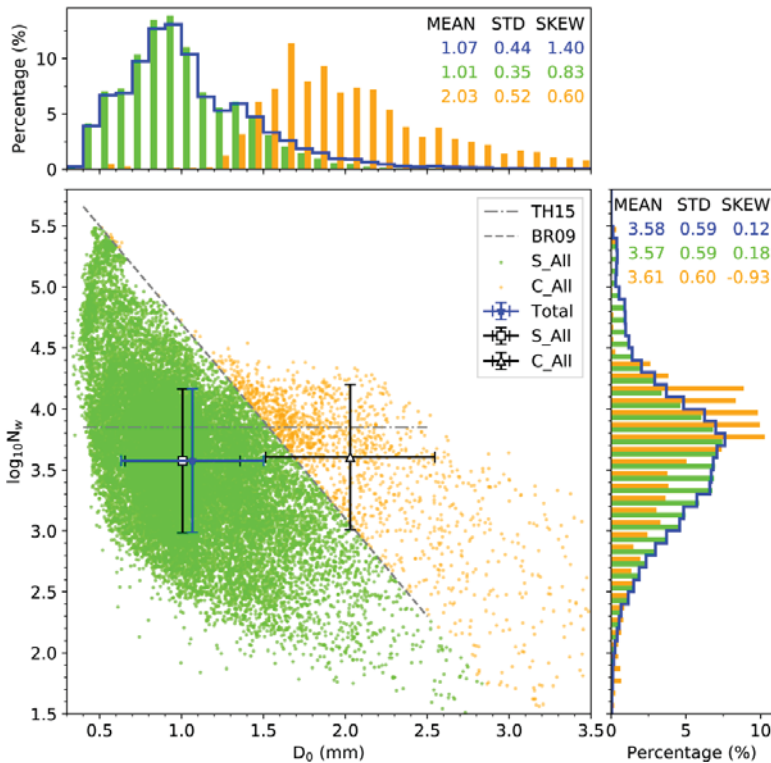
Figure 4. Scatterplot of $\log_{10} N_w$ vs. D_0 for stratiform (S_All, lime) and convective (C_All, orange) rain in the bottom left panel, as well as the corresponding relative frequency histograms in the top and bottom right panels. The unit of N_w is m^{-3} mm^{-1}. Rain types were classified by BR09 scheme. The C_All (S_All) dataset equals to the dataset of C_BR09 (S_BR09) denoted in Table 2. Blue curves in each histogram indicate the relative frequency of the entire dataset for $\log_{10} N_w$ and D_0. The mean (MEAN), standard deviation (STD) and skewness (SKEW) for the entire dataset, stratiform rain and convective rain are shown in colors in each histogram panel, whereas the MEAN values of $\log_{10} N_w$ vs. D_0 together with the respective $\pm 1 \times$ STD values are plotted as error bars. The dashed and dot-dashed grey lines represent the C−S separation lines of BR09 and TH15, respectively.

The normalized frequency of DSD sample occurrence is shown in Figure 5. Note that the TH15 line in W–D_0 domain (Figure 5b) can be generated by combining Equation (7) and (14). The highest frequency of occurrence is in the ranges of D_0 about 0.8–1.1 mm and $\log_{10} N_w$ about 3.2–4.1, corresponding to rainwater content W within 0.02–0.11 g m^{-3}. The distribution of normalized frequency of DSD in both $\log_{10} N_w$–D_0 and W–D_0 domains are similar to the analyses in Dolan et al. [46] (their Figure 2b,e) in the midlatitudes. Therefore, this study provides new evidence from midlatitude Asian (northern China) to further support such analysis.
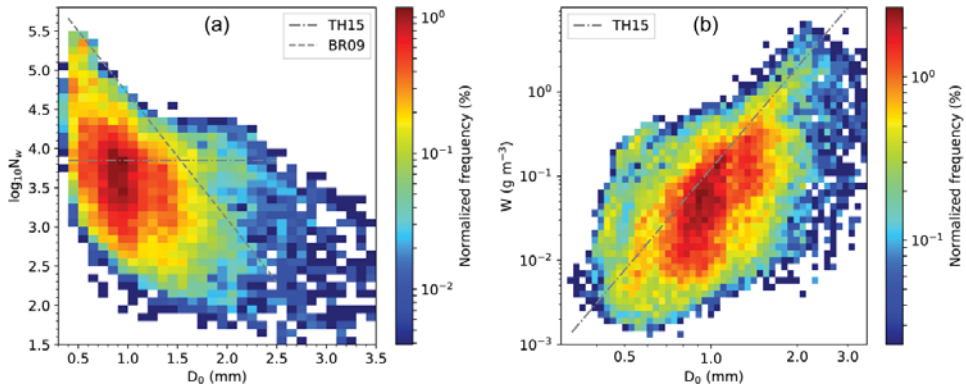
Figure 5. Normalized occurrence frequency of DSD sample in (**a**) $\log_{10} N_w - D_0$ and (**b**) $W - D_0$ domains. The dashed and dot-dashed lines represent the C−S separation lines from BR09 and TH15, respectively.

In Figure 6, the $\log_{10} N_w - D_0$ pairs are color coded by rain rate R and Z_H to investigate the interrelations among them. Similar patterns can be found in Figure 6a,b that the increases of both R and Z_H are proportional to the increases of $\log_{10} N_w$ and D_0, illustrating the internal relation between rain rate and radar reflectivity, or the Z_h–R relationship that will be discussed in Section 4. The TH15 line crosses all levels of R and Z_H, whereas BR09 line is almost equivalent to a threshold of R (8.6 mm h^{-1}) or Z_H (36.8 dBZ). Similar conclusion has been drawn for tropical, maritime regions with R = 10 mm h^{-1} and Z_H = 40 dBZ [53], which are slightly higher than our results.
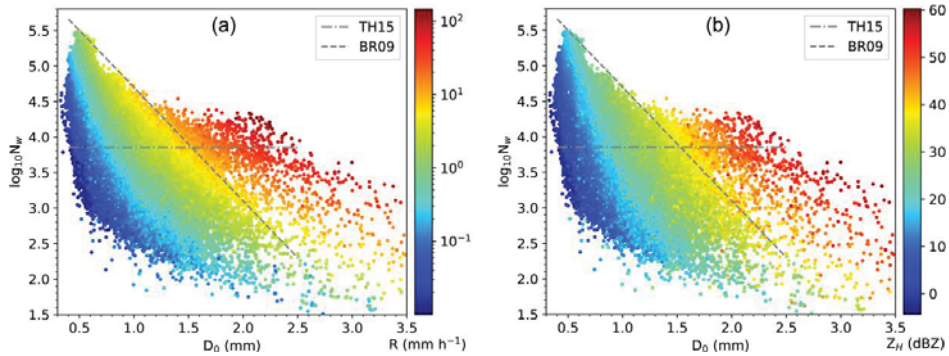
Figure 6. Scatterplots of $\log_{10} N_w$ vs. D_0 color coded by (**a**) R and (**b**) Z_H. The units of R and Z_H are in mm h^{-1} and dBZ, respectively. The dashed and dot-dashed lines represent the C−S separation lines from BR09 and TH15, respectively.

Interestingly, fewer DSD samples fell within $\log_{10} N_w > 4$ and $D_0 > 1$ mm (see Figures 4–6) compared to the results observed during the Asian Summer Monsoon Season in Eastrn [14] (their Figure 6) or Southern China [54] (their Figure 6), and in tropical, oceanic islands [53] (their Figure 14a,b). In addition, more DSD samples exist in the range above BR09 line but below TH15 line. Referring to Dolan et al. [46] and Bringi et al. [51], warm rain with the collision-coalescence process has a great contribution to the precipitation in Eastern and Southern China during the Asian Summer Monsoon Season and tropical, oceanic regions. On the contrary, mixed phase precipitaiton processes may dominante the rainfall microphysics near the disdrometer site in Beijing. The enhanced mixed phase precipitation processes can produce larger raindrops when the ice-based hydrometers melt, which need to be further investigated in future.

Datasets for convective and stratiform rain are further divided into months, as shown in the $\log_{10} N_w$–D_0 domain in Figure 7, to see the monthly variations in DSD and better compare with previous findings. For stratiform rain, the MEAN values of $\log_{10} N_w$ and D_0 in each month are all concentrated near the highest frequency of occurrences (Figure 5a), which corresponds to the "ambiguous" area in Figure 12 shown in reference [46]. For convective rain, those values are distributed in a larger range from the mixed area to the ice-based area (from April to August), as well as aggregation/riming area (September and October) in Figure 12 from Dolan et al. [46]. Note that for convective rain the MEAN values of $\log_{10} N_w$–D_0 pairs in months from May to August are almost all around the value of 3.61 and 2.03 mm for C_All dataset with minor variations. Their STD values are also similar, which means similar microphysical processes dominated the precipitation during these months. However, such characteristics are not observed in other months. Relatively larger $\log_{10} N_w$ and smaller D_0 indicate relatively more warm rain processes in April, while in September and October obviously lower $\log_{10} N_w$ and larger D_0 indicate the relatively intense ice-based processes, such as aggregation and riming that sharply exhausting the number of small size hydrometers but slowly increasing the size of drops. Such analyses demonstrate the seasonal variation of dominating microphysical processes in Beijing. Overall, all MEAN values for both rain types in each month are below the TH15 line, illustrating that different microphysical processes are dominating the precipitation between midlatitude and Eastern and Southern China during the Asian Summer Monsoon Season, as well as tropical, oceanic regions.
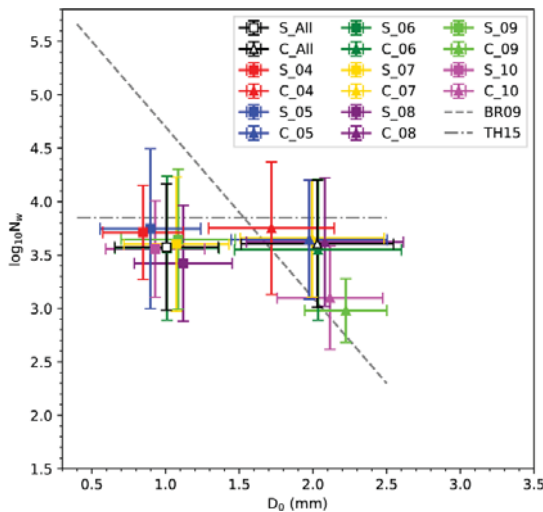
Figure 7. The MEAN values of $\log_{10} N_w$ vs. D_0 together with the respective $\pm 1 \times$ STD values plotted as error bars for convective (triangle) and stratiform (square) rain. The dataset for both rain types, including all data, are plotted in black, whereas the monthly results are indicated by different colors. The dashed and dot-dashed lines represent the C−S separation lines from BR09 and TH15, respectively.

3.3. Discussion on C−S Classification Schemes

The classification of precipitation into convective and stratiform is important in this study. Previous studies have proved that BR09 and TH15 schemes in $\log_{10} N_w$–D_0 domain are applicable based on the measurements not only from disdrometers but also from polarimetric radars [46,50,53,55,56]. As such, these classification approaches are adopted. However, there are also a few other C−S classification schemes. In order to reveal the impacts of the classification approach on the analysis results, this study also applied the C−S classification schemes described in Testud et al. [52] (hereafter referred as to TE01) and Bringi et al. [51] (hereafter referred as to BR03) for comparison purpose. Both schemes are

popularly used as well, and both are based on the variation of R with time and utilize 10 (5) adjacent DSD measurements at a 1-min (2-min) interval. The major difference between them is that TE01 assesses the values of R with an upper limit of 10 mm h^{-1} for stratiform rain, whereas BR03 evaluates the standard deviation of R (σ_R) with a lower threshold of 5 mm h^{-1} for convective rain. It should be mentioned that some DSDs may satisfy the conditions $R < 5$ mm h^{-1} and $\sigma_R \leq 1.5$ mm h^{-1} according to BR03, and, thus, fail to be classified as either stratiform or convective rain.

TH15 scheme is not suitable for Beijing, because no obvious peak of sample occurrences above Equation (14) can be found in Figure 5. Therefore, only integral rainfall parameters derived from BR09, BR03, and TE01 are listed in Table 2. Compared with BR09, both TE01 and BR03 schemes classify more convective (less stratiform) DSDs, which result in more (less) rainfall amount and a higher proportion of convective (stratiform) rain. However, almost all DSD parameter values for both rain types derived by TE01 and BR03 are not higher than those derived based on BR09, except the $\log_{10} N_w$ value for convective rain. Compared with Figure 4, convective rain classified by TE01 (Figure A1) and BR03 (Figure A2) in $\log_{10} N_w$–D_0 domain contain much more samples under BR09 line but above TH15 line, corresponding to the DSDs with higher number concentration but smaller size. As a result, the smallest $\log_{10} N_w$ but highest D_0 for convective rain are obtained by BR09.

For stratiform rain, the DSD parameters from S_TE01 are higher than those from S_BR03. For convective rain, however, it is the opposite (Table 2). Further study shows that the percentage of samples with $R > 5$ mm h^{-1} in C_BR03 is higher than that in C_TE01. In other words, the lower threshold of 5 mm h^{-1} for convective rain set in BR03 scheme plays a key role in the different results between TE01 and BR03.

In summary, for stratiform rain, the impacts of different C−S classification schemes are not distinct relative to convective rain, due to the higher number of samples for the former than the latter. Although DSDs classified by the aforementioned three schemes in $\log_{10} N_w$–D_0 domain can be separated by BR09 line in general (Figures 4, A1 and A2), the specific properties of DSDs could be different. The BR09 scheme is recommended, since it has been proved with radar observations [55,56].

4. Radar-Based Quantitative Precipitation Estimation

This study first computed Z_h and R using Equations (5) and (10), based on the DSD measurements, to support weather radar applications in Beijing. The power-law relation $Z_h = aR^b$ was then derived using nonlinear regression approach. It is well known that the Z_h–R relationship is dependent on local DSD variability, which can be influenced by many factors, such as rainfall type, climate regime, and orographic effect [17,35,57]. Finding a suitable Z_h–R relation for Beijing is also critical to RMAPS model for QPE forecast [36].

Figure 8 shows a scatterplot of Z_h–R pairs for both rain types classified by BR09 scheme along with the corresponding fitted power-law curves and equations. The fitted curve for the entire dataset is highlighted in black dots. For comparison, other four commonly used Z_h–R relationships are also indicated in Figure 8, including those for the continental stratiform rain ($Z_h = 200R^{1.6}$) [58], tropical systems ($Z_h = 250R^{1.2}$) [59], operational WSR-88D radars ($Z_h = 300R^{1.4}$) [60], and Meiyu convective rain in China ($Z_h = 368R^{1.21}$) [11]. Obviously, Z_h is proportional to R in the double logarithmic domain. Based on the fitted relations for the two rain types, for a given Z_h, higher R can be obtained using the stratiform relation than a convective algorithm. The relationship for the entire dataset (i.e., $Z_h = 265.14R^{1.399}$) is closer to the relationship for stratiform rain.
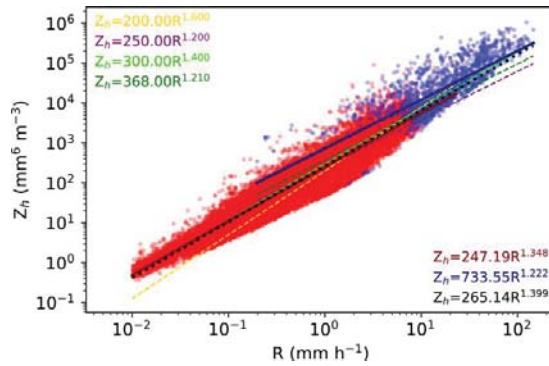
Figure 8. Scatterplot of Z_h (mm^6 m^{-3}) vs. R (mm h^{-1}) computed from PARSIVEL^2 DSD measurements for stratiform (red dots) and convective (blue dots) rain classified using BR09 scheme. The fitted power-law curves for stratiform and convective rain, as well as the entire dataset, are indicated by thick solid dark-red, solid dark-blue, and black dotted lines, respectively. The relationships for continental stratiform rain, $Z_h = 200R^{1.6}$ [58], tropical systems, $Z_h = 250R^{1.2}$ [59], the operational WSR-88D, $Z_h = 300R^{1.4}$ [60], and Meiyu convective rain, $Z_h = 368R^{1.21}$ [11] are also indicated in thin dashed yellow, purple, lime and green lines, respectively. Equations are overlaid using the same color with the corresponding curves.

It is worth noting that the relationship for the operational WSR-88D (thin dashed lime line) [60] is very similar to our result based on the entire dataset, which implies that the relationship $Z_h = 300R^{1.4}$ could potentially be employed for QPE in Beijing. For convective rain, both $Z_h = 250R^{1.2}$ and $Z_h = 368R^{1.21}$ will underestimate the rainfall intensities, likely due to the smaller diameter and higher number concentration of raindrops in these two climate regions than in Beijing (as detailed in Section 3.2). Compared with $Z_h = 300R^{1.4}$, $Z_h = 200R^{1.6}$ has relatively larger discrepancy compared to our result.

Although a suitable Z_h–R relationship can be helpful to retrieve rain rate from radar reflectivity, the dispersion of samples in Z_h–R domain is still large. For example, for a given $Z_h = 10^3$ mm^6 m^{-3}, R can range from 0.5–10 mm h^{-1} (Figure 8). To further investigate the essence of Z_h–R relationships from a microphysical point of view, the scatter distribution of Z_h–R pairs are color coded by D_0 and $\log_{10} N_w$ in Figure 9a,b. It is concluded that DSDs can be further grouped in size or number concentration in Z_h–R domain, which means the QPE could be further improved when considering more physical observables.
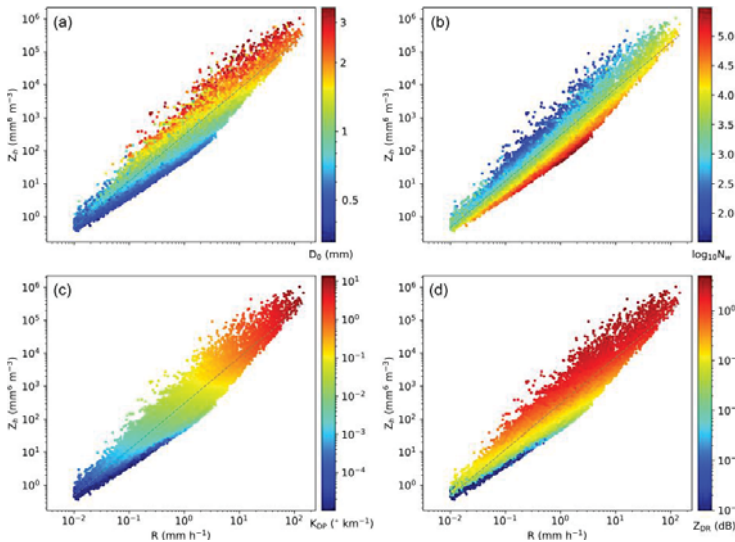
Figure 9. Scatterplots of Z_h (mm^6 m^{-3}) vs. R (mm h^{-1}) color coded by (**a**) D_0, (**b**) $\log_{10} N_w$, (**c**) K_{DP}, and (**d**) Z_{DR}. The $Z_h = 300R^{1.4}$ dashed line is superimposed for reference.

In addition, dual-polarization radar variables are computed using the *T*-matrix method. The polarimetric measurements are proven to be capable of improving the performance of QPE. Figure 9c,d show the distribution of Z_h versus R, color coded by K_{DP} and Z_{DR}, respectively. Overall, similar variation patterns can be seen compared with Figure 9a,b. This is not surprising, since D_0 and $\log_{10} N_w$ can essentially be derived from the combination of Z_h, Z_{DR}, and K_{DP} [34,45,61].

The distributions of Z_H, Z_{DR}, and K_{DP} are illustrated in Figure 10. It should be noted again that Z_H in dBZ is used in Figure 10a, while QPE estimators are fitted using Z_h in linear scale. The details of boxplot in the center of each panel are listed in Table 3. The median value of Z_H is about 20 dBZ, and the number of Z_H higher than 40 dBZ is less than 5%. A large amount of K_{DP} are smaller than 0.1 ° km^{-1}. The distribution of each parameter has two peaks: The first peak of Z_H and K_{DP} is close to their median values, while the second peaks are at about 27.5 dBZ and 0.07 ° km^{-1}, respectively. The two peaks of Z_{DR} are about 0.13 and 0.45 dB, and the median value lies between the two peaks.
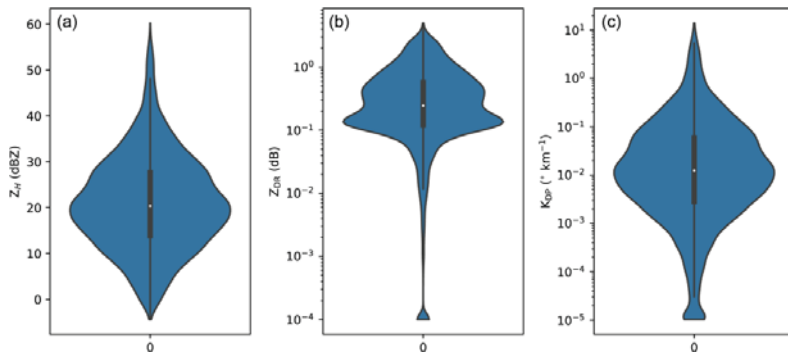
Figure 10. The distributions of (**a**) Z_H, (**b**) Z_{DR}, and (**c**) K_{DP} derived from DSD measurements using the *T*-matrix scattering approach.

Table 3. The quantiles of polarization radar variables derived from DSDs using the *T*-matrix scattering method.

	min	5%	25%	median	75%	95%	max
Z_H (dBZ)	-4.37	4.59	13.88	20.35	27.58	38.93	60.21
Z_{DR} (dB)	1.00×10^{-4}	1.95×10^{-2}	0.12	0.25	0.57	1.71	4.95
K_{DP} (° km^{-1})	1.03×10^{-5}	1.42×10^{-4}	2.85×10^{-3}	1.24×10^{-2}	5.82×10^{-2}	0.48	13.82

This study also derived the polarimetric radar rainfall relations $R_{dpr}(Z_h, Z_{DR})$, $R_{dpr}(K_{DP}, Z_{DR})$, and $R_{dpr}(K_{DP})$ using the least-squares method and compared with the Z_h–R relationships. Here, the subscript "dpr" represents Dual-Polarization Radar for short. The obtained estimators based on the total DSD dataset are listed as follows:

$$R_{dpr}(Z_h, Z_{DR}) = \alpha Z_h^{\beta} 10^{\gamma Z_{DR}}, \tag{15}$$

$$R_{dpr}(K_{DP}, Z_{DR}) = \alpha K_{DP}^{\beta} 10^{\gamma Z_{DR}}, \tag{16}$$

$$R_{dpr}(K_{DP}) = \alpha K_{DP}^{\beta}, \tag{17}$$

$$R_{dpr}(Z_h) = \alpha Z_h^{\beta}, \tag{18}$$

where α, β, and γ are generic coefficients and exponents in each relation. The specific values are listed in Table 4.

Table 4. The fitted parameters of radar QPE estimators (Equations (15)–(18)) derived using the total DSD dataset.

Parameters	$R_{dpr}(Z_h, Z_{DR})$	$R_{dpr}(K_{DP}, Z_{DR})$	$R_{dpr}(K_{DP})$	$R_{dpr}(Z_h)$
α	5.696×10^{-3}	23.045	15.375	6.986×10^{-2}
β	0.986	0.947	0.836	0.540
γ	-0.464	-0.101	——	——

In order to evaluate the application performance of various QPE estimators, the hourly rainfall amount (mm) derived using each radar rainfall relation is compared with collocated rain gauge observations (distance between disdrometer and gauge is less than 10 m). Figure 11a–d shows the scatter plots of rainfall estimated using radar relations versus gauge measurements. In addition, a set of evaluation metrics, including the Pearson correlation coefficient (PCC), standard deviation (STD), normalized mean absolute error (NMAE), and root-mean-square error (RMSE) are computed and indicated in Figure 11.

Obviously, $R_{dpr}(Z_h, Z_{DR})$ performs the best in terms of all evaluation metrics, followed by $R_{dpr}(K_{DP}, Z_{DR})$, $R_{dpr}(K_{DP})$, and then $R_{dpr}(Z_h)$. The estimated hourly rainfall amount from $R_{dpr}(Z_h, Z_{DR})$ (Figure 11a) is the closest to rain gauge measurements at low intensities. However, $R_{dpr}(K_{DP}, Z_{DR})$ provides the best estimation at higher rainfall intensities, especially during severe precipitation hours.

Recent studies [5,6] demonstrated that the combination of different estimators may improve the accuracy of QPE. However, their achievements were mainly based on S-band radar measurements. In this study, we attempted to extend this strategy to X-band applications. Similar thresholds to the Dual-Polarization Radar Operational Processing System version 2 (DROPS2) [5] are used at X-band: $Z_H = 37$ dBZ, $Z_{DR} = 0.185$ dB, and $K_{DP} = 0.03$ ° km^{-1}. For clarification, this paper referred to the implemented DROPS2.0 architecture as R_{dpr}(DROPS2–X). As expected, R_{dpr}(DROPS2–X) (Figure 11e) provides the best results among various rainfall relations, which demonstrates the feasibility of the thresholds applied on X-band dual-polarization radar variables. Compared with Figure 11b, R_{dpr}(DROPS2–X) inherits the advantage of $R_{dpr}(K_{DP}, Z_{DR})$ for all severe precipitation hours.

Nevertheless, it should be noted that except $R_{dpr}(Z_h)$, the differences among all other QPE estimators are not distinct: All have PCC higher than 0.98, STD and RMSE smaller than 1.0, and NMAE smaller than 0.2.
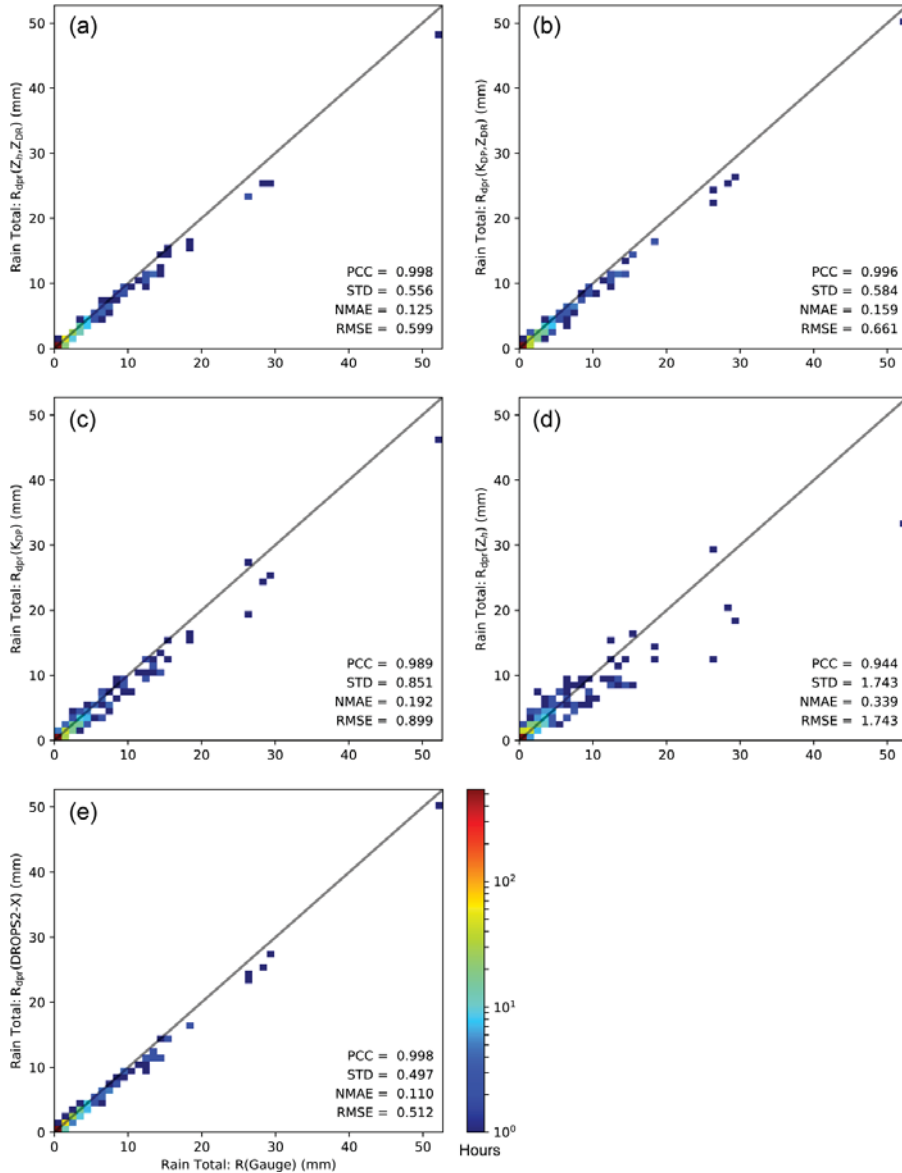
Figure 11. Scattergram (based on the total rainfall observations) of hourly rainfall estimates (mm) from various radar rainfall relations vs. rain gauge measurements: (**a**) $R_{dpr}(Z_h, Z_{DR})$, (**b**) $R_{dpr}(K_{DP}, Z_{DR})$, (**c**) $R_{dpr}(K_{DP})$, (**d**) $R_{dpr}(Z_h)$, and (**e**) $R_{dpr}(DROPS2-X)$. The grey diagonal straight line in each panel represents the 1–1 relationship. The quantitative evaluation results are also indicated in each panel, including the Pearson correlation coefficient (PCC), standard deviation (STD—mm), normalized mean absolute error (NMAE), and root-mean-square error (RMSE—mm).

5. Conclusions

To investigate the microphysical properties of surface precipitation and improve the accuracy of radar QPE, 14-month continuous PARSIVEL^2 measurements during 2017–2018 in Beijing, China, were analyzed in this study. After quality control, a total of 25,499 1-min DSD spectra were obtained, corresponding to 1013.78 mm of total rainfall. The major rainy periods were from June to August, which contributed to 81.5% of rainfall amount and 78.3% of total DSD samples. The least contribution of rainfall was from September. In October, the precipitation tends to be steady with relatively long time but low intensity.

DSD dataset was classified as stratiform and convective rain types using the BR09 C−S scheme [50] in $\log_{10} N_w$–D_0 domain. A large number of samples were identified as stratiform, which accounted for less than half of the total rainfall amount. The mean integral rainfall parameters, such as $\langle R \rangle$, $\langle \log_{10} N_w \rangle$, $\langle D_0 \rangle$, and three X-band dual-polarization variables, were higher in convective rain than stratiform rain. The occurrence of DSDs concentrated with D_0 and $\log_{10} N_w$ in the ranges of 0.8–1.1 mm and 3.2–4.1, respectively, which corresponds to W about 0.02–0.11 g m^{-3}. The increases of R and Z_H were proportional to the increases of $\log_{10} N_w$ and D_0, and BR09 line was equivalent to $R = 8.6$ mm h^{-1} and $Z_H = 36.8$ dBZ. The comparation with other C−S classification schemes showed the similar distribution in $\log_{10} N_w$–D_0 domain, but the detailed characteristics of DSDs among different schemes were different, with larger discrepancies in convective rain than stratiform rain. The different predominant microphysical processes in Beijing and other climate regions result in different DSD distributions in $\log_{10} N_w$–D_0 domain, especially for convective rain. Compared to the warm rain characterized by a collision-coalescence process in Eastern and Southern China during the Asian Summer Monsoon Season, as well as in tropical, oceanic regions, the precipitation in Beijing is dominated more by mixed phase precipitation microphysical processes. The melting large ice-phase hydrometers increased D_0 but decreased N_w compared to other climate regions. For stratiform rain, the mean values of $\log_{10} N_w$ and D_0 correspond to the high occurance ranges. For convective rain, three groups were separated, which showed distinct seasonal variations. The mean values of $\log_{10} N_w$–D_0 pairs from May to August (Group 1) clustered together while those from April (Group 2) and September-October (Group 3) were distributed on the two sides of Group 1 above the BR09 line. Group 2 tends to contain more warm rain processes, while Group 3 was dominated by intense ice-based processes, such as aggregation and riming that sharply decrease the number of small size hydrometers but slowly increase the particle size. This finding provides additional insight to precipitation microphysics in midlatitude Asian (northern China) and further appends the archievements of Dolan et al. [46].

In addition, dual-polarization radar variables were computed from the DSD dataset using the T-matrix scattering method and the radar-based QPE estimators were derived through nonlinear regression analysis. The estimated rainfall products using radar rainfall relations were also independently verified using collocated rain gauge measurements. It was concluded that for single-polarization variable, the fitted Z_h–R relationship, $Z_h = 265.14 R^{1.399}$, was almost coincident with the operational WSR-88D rainfall estimator [60], $Z_h = 300 R^{1.4}$; for dual-polarization radar applications, $R_{\mathrm{dpr}}(Z_h, Z_{\mathrm{DR}})$ performed the best for hourly rainfall estimation, while $R_{\mathrm{dpr}}(K_{\mathrm{DP}}, Z_{\mathrm{DR}})$ performed the best at high rainfall intensities. In addition, a blended algorithm is derived based on the architecture of DROPS2 [5] to enhance radar rainfall estimation. It was shown that $R_{\mathrm{dpr}}(\mathrm{DROPS2}\text{–}X)$ performed better than any individual QPE estimators at hourly scale. Future work will focus on the large scale application of $R_{\mathrm{dpr}}(\mathrm{DROPS2}\text{–}X)$ for the X-band dual-polarization radar network being deployed in Beijing.
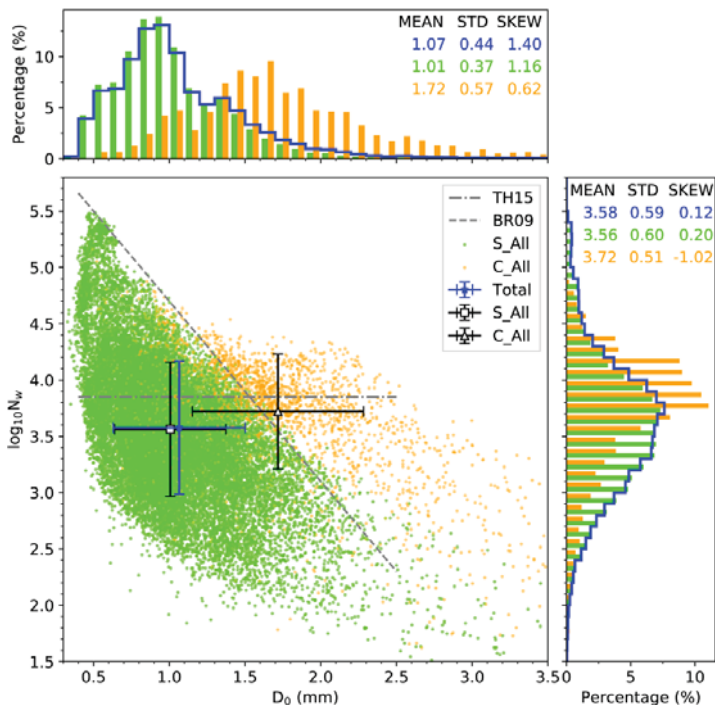
Appendix A

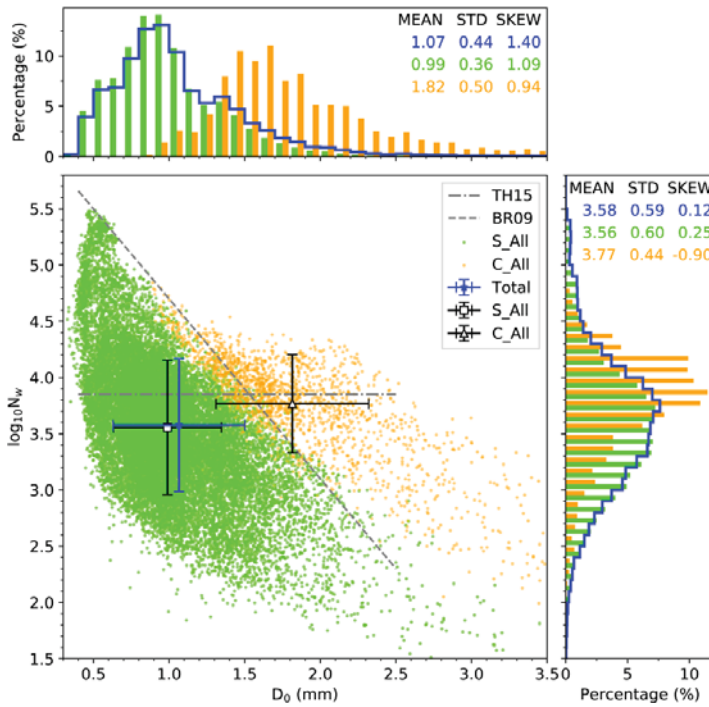**Figure A1.** As in Figure 4, but for the TE01 classification scheme.

Figure A2. As in Figure 4, but for BR03 classification scheme.

References

1. Houze, R.A.J. *Cloud Dynamics*, 1st ed.; California Academic Press: San Jose, CA, USA, 1993; p. 573.
2. Milbrandt, J.A.; Yau, M.K.; Multimoment, A. Bulk microphysics parameterization. Part i: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.* **2005**, *62*, 3051–3064. [CrossRef]
3. Zhang, G.; Sun, J.; Brandes, E. Improving parameterization of rain microphysics with disdrometer and radar observations. *J. Atmos. Sci.* **2006**, *63*, 1273–1290. [CrossRef]
4. Morrison, H.; Milbrandt, J.A. Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part i: Scheme description and idealized tests. *J. Atmos. Sci.* **2015**, *72*, 287–311. [CrossRef]
5. Chen, H.; Chandrasekar, V.; Bechini, R. An improved dual-polarization radar rainfall algorithm (DROPS2.0): Application in NASA ifloods field campaign. *J. Hydrometeorol.* **2017**, *18*, 917–937. [CrossRef]
6. Cifelli, R.; Chandrasekar, V.; Lim, S.; Kennedy, P.C.; Wang, Y.; Rutledge, S.A. A new dual-polarization radar rainfall algorithm: Application in colorado precipitation events. *J. Atmos. Ocean. Technol.* **2011**, *28*, 352–364. [CrossRef]
7. Salles, C.; Poesen, J.; Sempere-Torres, D. Kinetic energy of rain and its functional relationship with intensity. *J. Hydrol.* **2002**, *257*, 256–270. [CrossRef]
8. USDA_Agricultural_Research_Service. *Science Documentation Revised Universal Soil Loss Equation*; Version 2; 2013; p. 355. Available online: https://www.ars.usda.gov/ARSUserFiles/60600505/RUSLE/RUSLE2_Science_Doc.pdf. (accessed on 7 June 2019).
9. Caracciolo, C.; Prodi, F.; Battaglia, A.; Porcu, F. Analysis of the moments and parameters of a gamma DSD to infer precipitation properties: A convective stratiform discrimination algorithm. *Atmos. Res.* **2006**, *80*, 165–186. [CrossRef]
10. Chen, B.; Hu, Z.; Liu, L.; Zhang, G. Raindrop size distribution measurements at 4500 m on the tibetan plateau during TIPEX-III. *J. Geophys. Res.* **2017**, *122*, 11092–11106.
11. Chen, B.; Yang, J.; Pu, J. Statistical characteristics of raindrop size distribution in the meiyu season observed in eastern China. *J. Meteorol. Soc. Jpn.* **2013**, *91*, 215–227. [CrossRef]

12. Nzeukou, A.; Sauvageot, H.; Delfin Ochou, A.; Cheikh, M.F.K. Raindrop size distribution and radar parameters at cape verde. *J. Appl. Meteorol.* **2004**, *43*, 90–105. [CrossRef]

13. Tokay, A.; Short, D.A. Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteorol.* **1996**, *35*, 355–371. [CrossRef]

14. Wen, L.; Zhao, K.; Zhang, G.; Xue, M.; Liu, B.Z.; Chen, X. Statistical characteristics of raindrop size distributions observed in east china during the asian summer monsoon season using 2-d video disdrometer and micro rain radar data. *J. Geophys. Res.* **2016**, *121*, 2265–2282. [CrossRef]

15. Yuter, S.E.; Kingsmill, D.E.; Nance, L.B.; Löfflermang, M. Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow. *J. Appl. Meteorol. Climatol.* **2006**, *45*, 1450–1464. [CrossRef]

16. Zhang, G.; Vivekanandan, J.; Brandes, E.; Meneghini, R.; Kozu, T. The shape-slope relation in observed gamma raindrop size distributions: Statistical error or useful information? *J. Atmos. Ocean. Technol.* **2003**, *20*, 1106–1119. [CrossRef]

17. Rosenfeld, D.; Ulbrich, C.W. Cloud Microphysical Properties, Processes, and Rainfall Estimation Opportunities. In *Radar and Atmospheric Science: A collection of Essays in Honor of David Atlas, Meteorological Monographs*; Wakimoto, R.M., Srivastava, R., Eds.; American Meteorological Society: Boston, MA, USA, 2003; pp. 237–258.

18. Testik, F.Y.; Gebremichael, M. *Rainfall: State of the Science*; American Geophysical Union: Washington, DC, USA, 2013; p. 287.

19. Niu, S.; Jia, X.; Sang, J.; Liu, X.; Lu, C.; Liu, Y. Distributions of raindrop sizes and fall velocities in a semiarid plateau climate: Convective versus stratiform rains. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 632–645. [CrossRef]

20. Friedrich, K.; Higgins, S.; Masters, F.J.; Lopez, C.R. Articulating and stationary PARSIVEL disdrometer measurements in conditions with strong winds and heavy rainfall. *J. Atmos. Ocean. Technol.* **2013**, *30*, 2063–2080. [CrossRef]

21. Zwiebel, J.; Van Baelen, J.; Anquetin, S.; Pointin, Y.; Boudevillain, B. Impacts of orography and rain intensity on rainfall structure. The case of the HyMeX IOP7a event. *Q. J. R. Meteorol. Soc.* **2016**, *142*, 310–319. [CrossRef]

22. May, P.T.; Bringi, V.N.; Thurai, M. Do we observe aerosol impacts on dsds in strongly forced tropical thunderstorms? *J. Atmos. Sci.* **2011**, *68*, 1902–1910. [CrossRef]

23. Li, J.; Yu, R.C.; Wang, J.J. Diurnal variations of summer precipitation in Beijing. *Chin. Sci. Bull.* **2008**, *53*, 1933–1936. [CrossRef]

24. Yang, P.; Guo, Y.; Hou, W.; Liu, W. Spatial and diurnal characteristics of summer rainfall over Beijing Municipality based on a high-density AWS dataset. *Int. J. Climatol.* **2012**, *33*, 2769–2780. [CrossRef]

25. Chen, M.; Wang, Y.; Gao, F.; Xiao, X. Diurnal variations in convective storm activity over contiguous North China during the warm season based on radar mosaic climatology. *J. Geophys. Res.* **2012**, *117*, D20115. [CrossRef]

26. Chen, M.; Wang, Y.; Gao, F.; Xiao, X. Diurnal evolution and distribution of warm-season convective storms in different prevailing wind regimes over contiguous North China. *J. Geophys. Res.* **2014**, *119*, 2742–2763. [CrossRef]

27. Löfflermang, M.; Joss, J. An optical disdrometer for measuring size and velocity of hydrometeors. *J. Atmos. Ocean. Technol.* **2000**, *17*, 130–139. [CrossRef]

28. Tang, Q.; Xiao, H.; Guo, C.; Feng, L. Characteristics of the raindrop size distributions and their retrieved polarimetric radar parameters in northern and southern China. *Atmos. Res.* **2014**, *135–136*, 59–75. [CrossRef]

29. Li, H.; Cui, X.P.; Zhang, D.L. A statistical analysis of hourly heavy rainfall events over the Beijing metropolitan region during the warm seasons of 2007–2014. *Int. J. Climatol.* **2017**, *37*, 4027–4042. [CrossRef]

30. Zhang, X.Y.; Wang, Y.Q.; Lin, W.L.; Zhang, Y.M.; Zhang, X.C.; Gong, S.; Zhao, P.; Yang, Y.Q.; Wang, J.Z.; Hou, Q. Changes of atmospheric composition and optical properties over beijing—2008 olympic monitoring campaign. *Bull. Am. Meteorol. Soc.* **2009**, *90*, 1633–1651. [CrossRef]

31. Sun, Y.L.; Wang, Z.F.; Du, W.; Zhang, Q.; Wang, Q.Q.; Fu, P.Q.; Pan, X.L.; Li, J.; Jayne, J.; Worsnop, D.R. Long-term real-time measurements of aerosol particle composition in Beijing, China: Seasonal variations, meteorological effects, and source analysis. *Atmos. Chem. Phys.* **2015**, *15*, 10149–10165. [CrossRef]

32. Fan, J.; Rosenfeld, D.; Zhang, Y.; Giangrande, S.E.; Li, Z.; Machado, L.A.T.; Martin, S.T.; Yang, Y.; Wang, J.; Artaxo, P.; et al. Substantial convection and precipitation enhancements by ultrafine aerosol particles. *Science* **2018**, *359*, 411–418.

33. Jia, X.; Ma, X.; Bi, K.; Chen, Y.; Tian, P.; Gao, Y.; Liu, X.; He, H. Distribution of particle size and fall velocities of winter precipitation in Beijing. *Acta Meteorol. Sin.* **2018**, *76*, 148–159. (In Chinese)

34. Brandes, E.; Zhang, G.; Vivekanandan, J. Comparison of polarimetric radar drop size distribution retrieval algorithms. *J. Atmos. Ocean. Technol.* **2004**, *21*, 584–598. [CrossRef]

35. Chandrasekar, V.; Meneghini, R.; Zawadzki, I. Global and Local Precipitation Measurements by Radar. In *Radar and Atmospheric Science: A Collection of Essays in Honor of David Atlas. Meteorological Monographs*; Wakimoto, R.M., Srivastava, R., Eds.; American Meteorological Society: Boston, MA, USA, 2003; pp. 215–236.

36. Xie, Y.; Fan, S.; Chen, M.; Shi, J.; Zhong, J.; Zhang, X. An assessment of satellite radiance data assimilation in RMAPS. *Remote Sens.* **2019**, *11*, 54. [CrossRef]

37. Operating Instructions Present Weather Sensor OTT Parsivel[2]. Available online: https://www.ott.com/download/operating-instructions-present-weather-sensor-ott-parsivel2-without-screen-heating/ (accessed on 6 April 2019).

38. Jaffrain, J.; Berne, A. Experimental quantification of the sampling uncertainty associated with measurements from PARSIVEL disdrometers. *J. Hydrometeorol.* **2011**, *12*, 352–370. [CrossRef]

39. Gunn, R.; Kinzer, G.D. The terminal fall velocity for water droplets in stagnant air. *J. Atmos. Sci.* **1949**, *6*, 243–248. [CrossRef]

40. Atlas, D.; Srivastava, R.C.; Sekhon, R.S. Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.* **1973**, *11*, 1–35. [CrossRef]

41. Foote, G.B.; Toit, P.S.D. Terminal velocity of raindrops aloft. *J. Appl. Meteor.* **1969**, *8*, 249–253. [CrossRef]

42. Chen, Y.; Liu, J.; Duan, S.; Su, D.; Lv, D. Application of X-Band dual polarization radar in precipitation estimation in summer of Beijing. *Clim. Environ. Res.* **2012**, *17*, 292–302. (In Chinese)

43. Ma, J.; Su, D.; Jin, Y.; Li, R. The impact of attenuation of X-band dual linear polarimetric radar on hail recognition. *Plateau Meteorol.* **2012**, *31*, 825–835. (In Chinese)

44. Waterman, P.C. Matrix formulation of electromagnetic scattering. *Proc. IEEE* **1965**, *53*, 805–812. [CrossRef]

45. Bringi, V.N.; Chandrasekar, V. *Polarimetric Doppler Weather Radar*; Cambridge University Press: Cambridge, UK, 2001; p. 664.

46. Dolan, B.; Fuchs, B.; Rutledge, S.A.; Barnes, E.A.; Thompson, E.J. Primary modes of global drop size distributions. *J. Atmos. Sci.* **2018**, *75*, 1453–1476. [CrossRef]

47. Willis, P.T. Functional fits to some observed drop size distributions and parameterization of rain. *J. Atmos. Sci.* **1984**, *41*, 1648–1661. [CrossRef]

48. 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]

49. Skofronick-Jackson, G.; Petersen, W.A.; Berg, W.; Kidd, C.; Stocker, E.F.; Kirschbaum, D.B.; Kakar, R.; Braun, S.A.; Huffman, G.J.; Iguchi, T.; et al. The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1679–1695. [CrossRef]

50. Bringi, V.N.; Williams, C.R.; Thurai, M.; May, P.T. Using dual-polarized radar and dual-frequency profiler for DSD characterization: A case study from Darwin, Australia. *J. Atmos. Ocean. Technol.* **2009**, *26*, 2107–2122. [CrossRef]

51. Bringi, V.N.; Chandrasekar, V.; Hubbert, J.; Gorgucci, E.; Randeu, W.L.; Schoenhuber, M. Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *J. Atmos. Sci.* **2003**, *60*, 354–365. [CrossRef]

52. Testud, J.; Oury, S.; Black, R.A.; Amayenc, P.; Dou, X. The concept of 'normalized' distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *J. Appl. Meteorol.* **2001**, *40*, 1118–1140. [CrossRef]

53. Thompson, E.J.; Rutledge, S.A.; Dolan, B.; Thurai, M. Drop size distributions and radar observations of convective and stratiform rain over the equatorial indian and west pacific oceans. *J. Atmos. Sci.* **2015**, *72*, 4091–4125. [CrossRef]

54. Zhang, A.; Hu, J.; Chen, S.; Hu, D.; Liang, Z.; Huang, C.; Xiao, L.; Min, C.; Li, H. Statistical characteristics of raindrop size distribution in the monsoon season observed in southern china. *Remote Sens.* **2019**, *11*, 432. [CrossRef]
55. Thurai, M.; Bringi, V.N.; May, P.T. CPOL radar-derived drop size distribution statistics of stratiform and convective rain for two regimes in Darwin, Australia. *J. Atmos. Ocean. Technol.* **2010**, *27*, 932–942. [CrossRef]
56. Thurai, M.; Gatlin, P.N.; Bringi, V.N. Separating stratiform and convective rain types based on the drop size distribution characteristics using 2D video disdrometer data. *Atmos. Res.* **2016**, *169*, 416–423. [CrossRef]
57. Steiner, M.; Smith, J.A.; Uijlenhoet, R. A Microphysical interpretation of radar reflectivity–rain rate relationships. *J. Atmos. Sci.* **2004**, *61*, 1114–1131. [CrossRef]
58. Marshall, J.S.; Palmer, W.M.K. The distribution of raindrops with size. *J. Meteorol.* **1948**, *5*, 165–166. [CrossRef]
59. Rosenfeld, D.; Wolff, D.B.; Atlas, D. General probability-matched relations between radar reflectivity and rain rate. *J. Appl. Meteorol. Climatol.* **1993**, *32*, 50–72. [CrossRef]
60. Fulton, R.A.; Breidenbach, J.P.; Seo, D.J.; Miller, D.A.; O'Bannon, T. The WSR-88D rainfall algorithm. *Weather Forecast.* **1998**, *13*, 377–395. [CrossRef]
61. Zhang, G. *Weather Radar Polarimetry*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2016; p. 304.