remote sensing

Volume 1

Remote Sensing of Precipitation

Edited by
Silas Michaelides

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

www.mdpi.com/journal/remotesensing

MDPI

Remote Sensing of Precipitation

Remote Sensing of Precipitation

Volume 1

Special Issue Editor

Silas Michaelides

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade

Special Issue Editor
Silas Michaelides
The Cyprus Institute
Cyprus

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Remote Sensing* (ISSN 2072-4292) from 2018 to 2019 (available at: https://www.mdpi.com/journal/remotesensing/special_issues/precipitation_rs)

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**, *Article Number*, Page Range.

Volume 1
ISBN 978-3-03921-285-9 (Pbk)
ISBN 978-3-03921-286-6 (PDF)

Volume 1-2
ISBN 978-3-03921-289-7 (Pbk)
ISBN 978-3-03921-290-3 (PDF)

Contents

About the Special Issue Editor

Silas Michaelides is currently an Adjunct Professor at the Cyprus Institute. He is a former Director of the Department of Meteorology of Cyprus, a position that he reached having served through all the scientific ranks of this governmental organization for more than 40 years. He holds a Ph.D. in Meteorology, an M.Sc. in Agricultural Meteorology, a Master's degree in Public Sector Management, and a B.Sc. in Mathematics. He has published 115 papers in peer-reviewed scientific journals, many of which address precipitation and related issues. He has also published a book on precipitation. He was awarded the International Research Award for Young Scientist by the World Meteorological Organization. He is a member 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.

remote sensing

MDPI

Editorial

Editorial for Special Issue "Remote Sensing of Precipitation"

Silas Michaelides

The Cyprus Institute, Nicosia 2121, Cyprus; s.michaelides@cyi.ac.cy

Received: 2 February 2019; Accepted: 12 February 2019; Published: 14 February 2019

Abstract: This Special Issue hosts papers on all aspects of remote sensing of precipitation, including applications that embrace the use of remote-sensing techniques of precipitation in tackling issues, such as precipitation estimations and retrievals, along with their methodologies and corresponding error assessment; precipitation modelling including validation, instrument comparison, and calibration; understanding of cloud and precipitation microphysical properties; precipitation downscaling; precipitation droplet size distribution; assimilation of remotely sensed precipitation into numerical weather prediction models; and measurement of precipitable water vapor.

Keywords: precipitation; TRMM; GPM; IMERG; weather radar; precipitable water vapor; precipitation retrieval; rain rate; QPE

1. Introduction

Precipitation is a key meteorological variable and one of the most important pillars in the global water and energy balances. In its various forms, precipitation comprises the primary source of freshwater, which is vital for the sustainability of almost all human activities, and the significance of this natural resource is fundamental in effectively managing applications ranging from irrigation to industrial and household usage. Precipitation is often associated with extreme weather events, such as floods, droughts, and landslides, which can have tremendous socio-economic impacts on the local scale but quite frequently on the regional scale. It is well documented that precipitation is one of the elements that is greatly affected by climate change.

The sufficiently accurate observation or estimation of precipitation has important theoretical but also practical significance. Indeed, on the one hand, the accurate and timely understanding of its characteristics at the local, regional, and global scales is vital for an insightful understanding of the mechanisms underlying the Earth's atmosphere–ocean complex system; on the other hand, the timely issuance of flood warnings, drought monitoring, and effective water resource management is crucial in some areas.

Precipitation observations made with rain gauges provide relatively accurate point-based measurements of precipitation. However, due to precipitation heterogeneity across a broad spectrum of spatiotemporal scales, rain gauge observations most often represent only the local conditions and can result in potential errors when interpolated to larger scales, especially in areas characterized by complex terrain.

Additionally, the spatial distribution of rain gauges is extremely uneven, with sparse gauges in remote areas, less developed regions, oceanic surfaces, and areas with complex terrain. Therefore, in situ rain gauge data usually cannot meet the requirements of applications that depend on high spatiotemporal resolution precipitation data, some of which have already mentioned above.

Remotely sensed precipitation data have the advantage of adequate temporal resolution and fine spatial resolution with a wide coverage, enabling accurate precipitation estimates in data-scarce or ungauged regions. Remote sensing of precipitation is pursued through a broad spectrum of

continuously enhanced and upgraded instrumentation, embracing sensors which can be satellite-borne, ground-based over land, ship-borne, and aerial. The subsequent processing of the remotely sensed precipitation data provides added value to this vast source of precipitation data, widening the horizons for a more effective usage. Indeed, the advances in the establishment of reliable remotely sensed precipitation datasets has provided an opportunity to reliably retrieve the spatiotemporal pattern of precipitation.

Through the scientific contributions contained in this volume, this Special Issue aims at exposing the scientific community to the current advances in many important areas of the remote sensing of precipitation. By presenting state-of-the-art technologies and methodologies regarding the remote sensing of precipitation, this Special Issue aspires to stimulate further research in this unceasingly expanding scientific discipline.

The following Section delivers a summary of all the thirty-nine articles published in the current special issue. The articles are presented in alphabetical order based on the first author's name.

2. Overview of Contributions

Anagnostou et al. [1] evaluate the advantages of using X-band polarimetric (XPOL) radar as a means to fill the coverage gaps and improve complex terrain precipitation estimation and associated hydrological applications based on a field experiment conducted in an area of Northeast Italian Alps characterized by large elevation differences. The corresponding rainfall estimates from two operational C-band weather radar observations are compared to the XPOL rainfall estimates for a near-range (10–35 km) mountainous basin (64 km^2). In situ rainfall observations from a dense rain gauge network and two disdrometers (one 2D-video and one Parsivel) are used for ground validation of the radar-rainfall estimates. Ten storm events over a period of two years are used to explore the differences between the locally deployed XPOL versus longer range operational radar-rainfall error statistics.

Utilizing reanalysis and high sensitivity W-band radar observations from CloudSat, Behrangi and Richardson [2] assess simulated high-latitude precipitation and its future changes under the RCP8.5 global warming scenario. A subset of models was selected based on the smallest discrepancy relative to CloudSat and ERA-I reanalysis using a combined ranking for bias and Root Mean Squared Error (RMSE). After accounting for uncertainties introduced by internal variability due to CloudSat's limited four-year day-night observation period, RMSE provides greater discrimination between the models than a typical mean state bias criterion. Over 1976–2005 to 2071–2100, colder months experience larger fractional modeled precipitation increases than warmer months, and the observation-constrained models generally report a larger response than the full ensemble.

The Global Precipitation Measurement (GPM) mission Core Observatory is equipped with a dual-frequency precipitation radar (DPR) with the capability of measuring precipitation simultaneously at frequencies of 13.6 GHz (Ku-band) and 35.5 GHz (Ka-band). Since the GPM-DPR cannot use information from polarization diversity, radar reflectivity factor is the most important parameter used in all retrievals. Biswas and Chandrasekar [3] quantitatively compare GPM's observations of reflectivity at dual-frequency and instantaneous rainfall products against dual-polarization ground-based Next-Generation Radars (NEXRAD) from the GPM Validation Network (VN), adopting a 3-D volume matching technique. The ground radars are located in the southeastern plains of the Unites States of America.

In their paper, Cánovas-García et al. [4] study the accuracy of three Quantitative Precipitation Estimates (QPEs) obtained from remote sensing or ground-based radars and the extent to which they could complement or even be an alternative to rain gauge readings in the Iberian Peninsula. The first QPE is the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Cloud Classification System, a satellite-based QPE. The second and the third are QPEs from a meteorological radar with Doppler capabilities that work in the C-band. Pixel-to-point comparisons are made between the values yielded by the QPEs and those obtained by two networks

of rain gauges. The results obtained indicate that all the QPEs were well below the rain gauge values in extreme rainfall time slots.

The paper by Cersosimo et al. [5] presents a geostatistical downscaling procedure to improve the spatial resolution of precipitation data. The kriging method with external drift is applied to surface rain intensity (SRI) data obtained through the Operative Precipitation Estimation at Microwave Frequencies (OPEMW), which is an algorithm for rain rate retrieval based on Advanced Microwave Sounding Units (AMSU) and Microwave Humidity Sounder (MHS) observations. SRI data have been downscaled from the coarse initial resolution of AMSU-B/MHS radiometers to the fine resolution of Spinning Enhanced Visible and InfraRed Imager (SEVIRI) flying on board the Meteosat Second Generation (MSG) satellite. Orographic variables, such as slope, aspect, and elevation, are used as auxiliary data in kriging with external drift together with observations from the Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager (MSG-SEVIRI) in the water vapor band and in thermal-infrared.

The performances of the latest released Integrated Multi-satellite Retrievals for GPM mission (IMERG) version 5 (IMERG v5) and the TRMM Multi-satellite Precipitation Analysis 3B42 version 7 (3B42 v7) are evaluated and compared by Chen et al. [6] at multiple temporal scales over a semi-humid to humid climate transition area (Huaihe River basin) from 2015 to 2017. The impacts of rainfall rate, latitude, and elevation on precipitation detection skills are also investigated. Results indicate that both satellite estimates show a high Pearson correlation coefficient (above 0.89)—with gauge observations—and an overestimation of precipitation at monthly and annual scales. The mean daily precipitation of IMERG v5 and 3B42 v7 displays a consistent spatial pattern, and both characterize the observed precipitation distribution well, but 3B42 v7 tends to markedly overestimate precipitation over water bodies.

Based on rain drop size distribution (DSD) measurements from 16 disdrometers located in Lausanne, Switzerland, Ghada et al. [7] present evidence that rain DSD differs among general weather patterns (GWLs). GWLs were successfully linked to significant variations in the rain microstructure characterized by the most important rain properties: rain intensity, mass weighted rain drop diameter, and rain drop concentration as well as parameters of the Z–R relation. The results highlight the potential to improve radar-based estimations of rain intensity, which is crucial for several hydrological and environmental applications.

Huang et al. [8] assess the performance of the latest version V5B of IMERG Early and Final Runs over southern China during six extremely heavy precipitation events. Observations from a dense network composed of 2449 rain gauges are used as a reference to quantify the performance in terms of spatiotemporal variability, probability distribution of precipitation rates, contingency scores, and bias analysis. The results show that both IMERG with gauge calibration and without gauge correction generally capture the spatial patterns of storm-accumulated precipitation with moderate to high correlation coefficients. Moreover, IMERG with gauge calibration and without gauge correction captured the area-average hourly series of precipitation over rainfall centers with a high correlation coefficient. Lastly, IMERG with gauge calibration tends to underestimate precipitation, especially the rainfall over the rainfall centers, when compared to IMERG without gauge correction.

Ivanov et al. [9] present a pre-processing approach adopted for the radar reflectivity data assimilation and results of simulations with the Harmonie numerical weather prediction model. The proposed method creates a 3-D regular grid in which a horizontal size of meshes coincides with the horizontal model resolution. This minimizes the representative error associated with the discrepancy between resolutions of informational sources. After such pre-processing, horizontal structure functions and their gradients for radar reflectivity maintain the sizes and shapes of precipitation patterns similar to those of the original data. The method shows an improvement of precipitation prediction within the radar location area in both the rain rates and spatial pattern presentation. It redistributes precipitable water with smoothed values over the common domain since the control runs show, among several sub-domains with increased and decreased values, correspondingly. It also reproduces the mesoscale belts and cell patterns of sizes from a few to ten kilometers in precipitation fields.

In the paper by Khan et al. [10], the performance of Level-3 gridded GPM-based precipitation products (IMERG) is assessed against two references over oceans: the OceanRAIN dataset, derived from oceanic shipboard disdrometers, and a satellite-based radar product (the Level-3 Dual-frequency Precipitation Radar, 3DPRD). Daily IMERG products (early, late, final) and microwave-only (MW) and Infrared-only (IR) precipitation components are evaluated at four different spatial resolutions (0.5○, 1○, 2○, and 3○) during a 3-year study period (March 2014–February 2017). Their performance is assessed based on both categorical and continuous performance metrics, including correlation coefficient, probability of detection, success ratio, bias, and RMSE. A triple collocation analysis (TCA) is also presented to further investigate the performance of these satellite-based products. Overall, the IMERG products show an underestimation with respect to OceanRAIN. Rain events in OceanRAIN are correctly detected by all IMERG products ~80% of the time. IR estimates show relatively large errors and low correlations with OceanRAIN compared to the other products. The MW component performs better than other products in terms of both categorical and continuous statistics. TCA reveals that 3DPRD performs consistently better than OceanRAIN in terms of RMSE and coefficient of determination at all spatial resolutions.

Dual-frequency Global Navigation Satellite Systems (GNSSs) enable the estimation of Zenith Tropospheric Delay (ZTD) which can be converted to Precipitable Water Vapor (PWV). The density of existing GNSS monitoring networks is insufficient to capture small-scale water vapor variations that are especially important for extreme weather forecasting. A densification with geodetic-grade dual-frequency receivers is not economically feasible. Cost-efficient single-frequency receivers offer a possible alternative. Krietemeyer et al. [11] study the feasibility of using low-cost receivers to increase the density of GNSS networks for retrieval of PWV. The authors process one year of GNSS data from an International GNSS Service station and two co-located single-frequency stations. Additionally, in another experiment, the Radio Frequency signal from a geodetic-grade dual-frequency antenna was split to a geodetic receiver and two low-cost receivers.

The paper by Le et al. [12] aims at exploring the capacity of the satellite-based rainfall product Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), including 3B42V7 research data and its real-time 3B42RT data, by comparing them against data from 29 ground observation stations over the lower part of the Red-Thai Binh River Basin from March 2000 to December 2016. Various statistical metrics were applied to evaluate the TMPA products. The results show that both 3B42V7 and 3B42RT had weak relationships with daily observations, but 3B42V7 data had strong agreement on the monthly scale compared to 3B42RT. Seasonal analysis shows that 3B42V7 and 3B42RT underestimate rainfall during the dry season and overestimate rainfall during the wet season, with high bias observed for 3B42RT. In addition, detection metrics demonstrate that TMPA products could detect rainfall events in the wet season much better than in the dry season. When rainfall intensity is analyzed, both 3B42V7 and 3B42RT overestimate the no rainfall event during the dry season but underestimate these events during the wet season.

The study by Lu et al. [13] assesses the performance of the latest GPM Integrated Multi-satellite Retrievals (IMERG V5) and Global Satellite Mapping of Precipitation version 7 (GSMaP V7) products and their hydrological application over the Tibetan Plateau. Two IMERG Final Run products (uncalibrated IMERG [IMERG-UC]) and gauge-calibrated IMERG [IMEEG-C]) and two GSMaP products (GSMaP Moving Vector with Kalman Filter [GSMaP-MVK] and gauge-adjusted GSMaP [GSMaP-Gauge]) were evaluated from April 2014 to March 2017. Results show that all four satellite precipitation products could generally capture the spatial patterns of precipitation over the Tibetan Plateau.

Martins Costa do Amaral et al. [14] apply the Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF) consolidated radar data processing to the X-band radar used in the CHUVA (Cloud Processes of the Main Precipitation Systems in Brazil: A Contribution to Cloud-Resolving Modeling and to the Global Precipitation Measurement) campaigns in Brazil. They apply the well-established H-SAF validation procedure to these data

and verify the quality of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) H-SAF operational passive microwave precipitation products in two regions of Brazil, namely, Vale do Paraíba and Manaus. These products are based on two rainfall retrieval algorithms: the physically based Bayesian Cloud Dynamics and Radiation Database (CDRD algorithm) for SSMI/S sensors and the Passive microwave Neural network Precipitation Retrieval algorithm (PNPR) for cross-track scanning radiometers (AMSU-A/AMSU-B/MHS sensors) and for the Advanced Technology Microwave Sounder (ATMS) sensor.

Remote sensing techniques provide extensive spatial coverage compared to ground-based rainfall data, but it is imperative to assess the quality of the estimates. Previous studies underline, at regional scale in the Amazon Basin in Brazil, and for some years, the efficiency of the Tropical Rainfall Measurement Mission (TRMM) 3B42 Version 7 (V7) (denoted by 3B42) daily product data to provide a good view of the rainfall time variability, which is important to understand the impacts of El Nino Southern Oscillation. The study by Michot et al. [15] aims at enhancing the knowledge about the quality of this product on the entire Amazon Basin and to provide a useful understanding about its capacity to reproduce the annual rainfall regimes. For that purpose, the authors compare 3B42 against 205 quality-controlled rain gauge measurements, for the period from March 1998 to July 2013, with the aim to reveal whether 3B42 is reliable for climate studies.

Hurricane Harvey, one of the most extreme events in recent history, advanced as a category IV storm and brought devastating rainfall to the region of Houston, Texas, the United States of America, during 25–29 August 2017. Omranian et al. [16] study the ability of the recent IMERG final-run product to capture the magnitudes and spatiotemporal patterns of rainfall resulting from hurricane Harvey. The results indicate that the performance of the IMERG product is satisfactory in detecting the spatial variability of the storm and it reconstructs precipitation with nearly 62% accuracy, although it systematically under-represents rainfall in coastal areas and over-represents rainfall over the high-intensity regions.

Rahman et al. [17] evaluate the performance of satellite-based precipitation products (SPPs), including GPM-based Integrated Multi-Satellite Retrievals for GPM (IMERG), TMPA (3B43-v7), PERSIANN and PERSIANN-CDR (Climate Data Record), over Pakistan based on Surface Precipitation Gauges (SPGs) at spatial and temporal scales. A novel ensemble precipitation (EP) algorithm is developed by selecting the two best SPPs using the paired sample t-test and principal component analysis (PCA). The SPPs and EP algorithm are evaluated over five climate zones (ranging from glacial Zone-A to hyper-arid Zone-E) based on six statistical metrics.

Ramsauer et al. [18] examine the performance of NASA's IMERG satellite precipitation dataset in capturing the spatio-temporal variability of weather events compared to the gauge-adjusted quality-controlled RADOLAN (Radar Online Adjustment) RW dataset from the German Weather Service. The results indicate significant differences between the datasets. Overall, GPM IMERG overestimates the quantity of precipitation compared to RADOLAN, especially in the winter season. Moreover, shortcomings in performance arise in this season with significant erroneous detections, yet precipitation events are also missed compared to the weather radar data. Additionally, along secondary mountain ranges and the Alps, topographically-induced precipitation is not represented in GPM data, which generally shows a lack of spatial variability in rainfall and snowfall estimates due to lower resolutions.

Rehman et al. [19] present an assessment on the validation and performance of both TMPA 3-hourly rainfall products, i.e., 3B42 near-real-time and 3B42 research version hourly rain rates, at various time scales for the period of 2007 to 2010 in Pakistan. The major objective of the study is to evaluate TMPA's rainfall rates at 3-h time intervals by using 10-min temporal resolution telemetric rain gauge (TRG) observations. The results show that the performance of TMPA research version in the detection of rain/no-rain events is quite good and has the best scores in a post-monsoon season. The performance of the TMPA near-real time product is inferior compared to the research version. Both TMPA products perform least in the winter season and have the lowest scores for performance

indices. Furthermore, the performance of TMPA products is questionable at light rain rates, but it improves at higher rain intensities.

Retalis et al. [20] assess the accuracy of the new IMERG satellite precipitation product over Cyprus on a monthly basis and during extreme events during the first years of GPM. This study is a first attempt to validate a specific GPM precipitation product by using in situ surface observations over the island of Cyprus. The satellite product was validated against the gauge data by simply comparing the former with the latter within each satellite cell. It is concluded that a very good agreement on monthly level was established, although IMERG tends to be underestimated as elevation goes higher, especially for areas higher than 1000 m elevation. Regarding daily evaluation, in terms of extreme recorded events, it seems that there exists a case dependency, while no specific correlation with elevation was established.

Ricciardelli et al. [21] investigate the value of satellite-based observational algorithms in supporting numerical weather prediction (NWP) for improving the alert and monitoring of extreme rainfall events. To this aim, the analysis of the very intense precipitation that affected the city of Livorno, Italy, on 9 and 10 September 2017 is performed by applying three remote-sensing techniques based on satellite observations at infrared/visible and microwave frequencies and by using maps of accumulated rainfall from the weather research and forecasting (WRF) model: The satellite-based observational algorithms are the precipitation evolving technique (PET); the rain class evaluation is from the infrared and visible observations (RainCEIV) technique; and the cloud classification mask is the coupling of statistical and physics methods (C-MACSP).

In view of the need to evaluate the accuracy of the new suite of GPM products in different regions of the globe, the study by Rozante et al. [22] aims at assessing the IMERG-Final and GSMaP-Gauge retrievals in order to replace the TMPA research version in several regions with different precipitation regimes in Brazil. In this respect, the performance of three satellite-based rainfall estimation products are evaluated in five Brazilian regions with different precipitation regimes, and the whole Brazilian territory, during the period from April 2014 to February 2017. This study shows that GSMaP-G and IMERG-F precipitation products exhibit better performance compared to the current TMPA-V7.

The paper by Rysman et al. [23] describes a new algorithm that is able to detect snowfall and retrieve the associated snow water path (SWP), for any surface type, using the GPM Microwave Imager (GMI). The algorithm is tuned and evaluated against coincident observations of the Cloud Profiling Radar (CPR) onboard CloudSat. It is composed of three modules for (i) snowfall detection, (ii) supercooled droplet detection, and (iii) SWP retrieval. This algorithm takes into account environmental conditions to retrieve SWP and does not rely on any surface classification scheme. Three case studies of snowfall events are investigated, and a 2-year high resolution snowfall occurrence distribution is presented.

Sanò et al. [24] describe a new rainfall rate retrieval algorithm, developed within the EUMETSAT H SAF program, based on the Passive microwave Neural Network Precipitation Retrieval approach (PNPR v3), designed to work with the conically scanning GMI. A new rain/no-rain classification scheme, also based on the Neural Network approach, which provides different rainfall masks for different minimum thresholds and degree of reliability, is also described. The algorithm is trained on an extremely large observational database built from GPM global observations between 2014 and 2016 where the NASA 2B-CMB (V04) rainfall rate product is used as a reference.

Satgé et al. [25] present an assessment of the potential benefits of the successive GPM-based SPEs product versions that include the Integrated Multi–Satellite Retrievals for GPM (IMERG) version 3 to 5 (–v03, –v04, –v05) and the Global Satellite Mapping of Precipitation (GSMaP) version 6 to 7 (–v06, –v07). The analysis is conducted over different geomorphic and meteorological regions of Pakistan while using 88 precipitation gauges as a reference. The results show a clear enhancement in precipitation estimates that were derived from the IMERG–v05 in comparison to its two previous versions IMERG–v03 and –v04; GSMaP–v07 precipitation estimates are more accurate than the previous GSMaP–v06.

In the study by Senent-Aparicio et al. [26], four widely used global precipitation datasets (GPDs) are evaluated: The Tropical Rainfall Measuring Mission (TRMM) 3B43, the Climate Forecast System Reanalysis (CFSR), PERSIANN, and the Multi-Source Weighted-Ensemble Precipitation (MSWEP) against point gauge and gridded dataset observations. Multiple monthly water balance models (MWBMs) are used in four different mesoscale basins that cover the main climatic zones of Peninsular Spain. Results underscore the superiority of the national gridded dataset, although TRMM provides satisfactory results in simulating streamflow.

The aim of the article by Sokol et al. [27] is to describe two new functionalities that they added to the radar data processing to study the cloud structures, which is their research purpose. Specifically, these authors deal with (i) the estimation of vertical air velocity and terminal velocity of hydrometeors and (ii) the classification of hydrometeors for which the vertical velocity and terminal velocity are the input parameters. Hydrometeors are considered to be any kind of liquid or solid water particles in the atmosphere that can result in precipitation which may or may not reach the ground in the form of graupel, rain, snow, or hail. The computational methods are illustrated with a thunderstorm event that occurred over Milešovka observatory, Czech Republic, on 1 June 2018.

Precise estimates of precipitation are required for many environmental tasks, including water resources management, improvement of numerical model outputs, nowcasting, and evaluation of anthropogenic impacts on global climate. Nonetheless, the availability of such estimates is hindered by technical limitations. Rain gauge and ground radar measurements are limited to land, and the retrieval of quantitative precipitation estimates from satellite has several problems including the indirectness of infrared-based geostationary estimates and the low orbit of those microwave instruments capable of providing a more precise measurement but suffering from poor temporal sampling. To overcome such problems, data fusion methods have been devised to take advantage of synergisms between available data, but these methods also present issues and limitations. Future improvements in satellite technology are likely to follow two strategies. One is to develop geostationary millimeter-submillimeter wave soundings, and the other is to deploy a constellation of improved polar microwave sensors. Tapiador et al. [28] compare both strategies using a simulated precipitation field. The results show that spatial correlation and RMSE would be little affected at the monthly scale in the constellation, but the precise location of the maximum of precipitation could be compromised.

Multiple scattering in the microwave frequencies has to be accounted for in precipitation retrieval algorithms. Tapiador et al. [29] study the effects of random arrangements of particles in space. In their study, such effects are described on the basis of modeled microwave scattering properties using the T-matrix formalism on the simplified case of spherical, non-intersecting raindrops. The results of measured rain drop size distributions (RDSDs) reveal that the random distribution of particles in space have a measurable but small effect on the scattering because of changes in the relative phases among the particles.

In the paper by Tapiador et al. [30], the oceanic precipitation projections of eight Regional Climate Models (RCMs) for the present-climate and the A2 SRES (Special Report on Emissions Scenarios) future-climate scenario have been compared with satellite and offshore gauge estimates. Notwithstanding observational errors and several other uncertainties, the results show that the RCMs consistently reproduce the observed oceanic precipitation of Europe, thus increasing the confidence in such models being capable of estimating changes in the future oceanic precipitation patterns. By integrating the uncertainties in both the observational and the modeled oceanic precipitation, a consensus estimate is made.

The principal objective of the study by Tsarpalis et al. [31] is to present and evaluate an advanced dust wet deposition scheme in the WRF model coupled with Chemistry (WRF-Chem). The integration of a dust wet deposition scheme into the WRF-Chem model is assessed through a case study of large-scale Saharan dust transport over the Eastern Mediterranean that is characterized by severe wet deposition over Greece. An acceptable agreement was found between the calculated and measured

near-surface PM_{10} concentrations as well as when model-estimated atmospheric optical depth (AOD) is validated against the AErosol RObotic NETwork (AERONET) measurements.

Varlas et al. [32] conducted a hydrometeorological analysis of a flash flood event that took place in the sub-urban area of Mandra, western Attica, Greece, using remote-sensing observations and the Chemical Hydrological Atmospheric Ocean Wave System (CHAOS) modeling system that includes the Advanced WRF (WRF-ARW) model and the hydrological model (WRF-Hydro). The flash flood was caused by a severe storm during the morning of 15 November 2017 around the Mandra area resulting in extensive damages and 24 fatalities. The X-band dual-polarization (XPOL) weather radar of the National Observatory of Athens (NOA), which was also used in this analysis, recorded precipitation rates reaching 140 mm/h in the core of the storm. The findings of this study demonstrate the potential benefit of using high-resolution observations from a locally deployed X-band dual-polarization radar as an additional forcing component in model precipitation simulations.

The near-real-time legacy product of Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis (3B42RT) and the equivalent products of Integrated Multi-satellite Retrievals for GPM (IMERG-E and IMERG-L) were evaluated and compared over mainland China from 1 January 2015 to 31 December 2016 at the daily timescale against rain gauge measurements by Wu et al. [33]. Results show that: (i) Both 3B42RT and IMERG products overestimate light rain, while underestimating moderate rain to heavy rainstorm; (ii) higher rainfall intensity associates with better detection; and (iii) both 3B42RT and IMERG products perform better in wet areas with relatively heavy rainfall intensity and/or during wet season than in dry areas with relatively light rainfall intensity and/or during dry season.

Wu et al. [34] assess the performance and hydrological utility of merged precipitation products at the current technical level of integration. A newly developed merged precipitation product, namely, the Multi-Source Weighted-Ensemble Precipitation (MSWEP) Version 2.1, is evaluated in this study based on rain gauge observations and the Variable Infiltration Capacity (VIC) model for the upper Huaihe River Basin, China. For comparison, three SPPs, including Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) Version 2.0, Climate Prediction Center MORPHing technique (CMORPH) bias-corrected product Version 1.0, and TMPA 3B42 Version 7, are appraised.

A case study on the impact of assimilating satellite radiance observation data into the rapid-refresh multi-scale analysis and prediction system (RMAPS) is presented by Xie et al. [35]. This case study targets the 48 h period from 19–20 July 2016 which was characterized by the passage of a low pressure system that produced heavy rainfall over North China. Two experiments were performed, and 24 h forecasts were produced every 3 h. The results indicate that the forecast prior to the satellite radiance data assimilation could not accurately predict heavy rainfall events over Beijing and the surrounding area. The assimilation of satellite radiance data from the advanced microwave sounding unit-A (AMSU-A) and microwave humidity sounding (MHS) improves the skills of the quantitative precipitation forecast to a certain extent.

Zhang et al. [36] investigate the accuracies of the estimated precipitation in the Tianshan Mountains, China, from three satellite products, namely, IMERG, TRMM 3B42, and CMORPH. These products are evaluated through comparisons with observations from 46 stations. The study period is from April 2014 to March 2015, which was the overlapping time period of the three respective satellite missions. The findings of this study suggest that GPM may outperform its predecessors in the mid- or high-latitude dryland areas but not in the tropical mountainous areas.

A rainfall retrieval algorithm for tropical cyclones (TCs) using 18.7 and 36.5 GHz of vertically and horizontally polarized brightness temperatures (Tbs) from the Microwave Radiation Imager (MWRI) is presented by Zhang et al. [37]. The beamfilling effect is corrected based on ratios of the retrieved liquid water absorption and theoretical Mie absorption coefficients. To assess the performance of this algorithm, MWRI measurements are matched with the National Snow and Ice Data Center (NSIDC) precipitation for six TCs. The comparison between MWRI and NSIDC rain rates is relatively

encouraging. A comparison of pixel-to-pixel retrievals shows that MWRI retrievals are constrained to reasonable levels for most rain categories.

Zhang et al. [38] evaluate the use of precipitation forecasts from a numerical weather prediction (NWP) model for near-real-time satellite precipitation adjustment based on 81 flood-inducing heavy precipitation events in seven mountainous regions over the conterminous United States of America. The study is facilitated by the National Center for Atmospheric Research (NCAR) real-time ensemble forecasts ("model"), the IMERG near-real-time precipitation product ("raw IMERG"), and the Stage IV multi-radar/multi-sensor precipitation product ("Stage IV") used as a reference. The authors evaluate four precipitation datasets (the model forecasts, raw IMERG, gauge-adjusted IMERG, and model-adjusted IMERG) through comparisons against Stage IV at six-hourly and event-length scales. The raw IMERG product consistently underestimates heavy precipitation in all study regions while the domain average rainfall magnitudes exhibited by the model are fairly accurate.

The performance of three SPPs over the upper catchment of the Red River Basin in China, for the time period of 1998–2010, is assessed by Zhang et al. [39]. The SPPs include TRMM 3B42 V7, CMORPH_CRT (CMORPH Bias-corrected product), and PERSIANN_CDR (Climate Data Record). The main objectives of this study are to (i) statistically evaluate the quality of the three SPPs through comparison with rain gauge observations and (ii) comprehensively explore and compare the capability of these three SPPs in streamflow simulations using GR (Génie Rural) hydrological models at daily and monthly scales.

3. Conclusions

The broad worldwide authorship in this Special Issue is a postulation that remote sensing of precipitation is a hot issue. Indeed, the scientific community invests much effort in revealing the power of tools available for remote sensing of this variable. In many cases, the supremacy of remote sensing over other traditional methodologies is demonstrated. However, many studies in this volume stress the significance of conventional measurements in validating remotely sensed estimations of precipitation.

Acknowledgments: The Guest Editor of this Special Issue would like to thank all authors who have contributed to this volume for sharing their scientific results and for their excellent collaboration. Special thanks are due to the community of distinguished reviewers for their valuable and insightful inputs. The Remote Sensing editorial team is gratefully acknowledged for its support during all phases of the endeavor to successfully complete this volume.

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

References

1. Anagnostou, M.N.; Nikolopoulos, E.I.; Kalogiros, J.; Anagnostou, E.N.; Marra, F.; Mair, E.; Bertoldi, G.; Tappeiner, U.; Borga, M. Advancing precipitation estimation and streamflow simulations in complex terrain with X-Band dual-polarization radar observations. *Remote Sens.* **2018**, *10*, 1258. [CrossRef]
2. Behrangi, A.; Richardson, M. Observed high-latitude precipitation amount and pattern and CMIP5 model projections. *Remote Sens.* **2018**, *10*, 1583. [CrossRef]
3. Biswas, S.; Chandrasekar, V. Cross-validation of observations between the GPM dual-frequency precipitation radar and ground based dual-polarization radars. *Remote Sens.* **2018**, *10*, 1773. [CrossRef]
4. Cánovas-García, F.; García-Galiano, S.; Alonso-Sarría, F. Assessment of satellite and radar quantitative precipitation estimates for real time monitoring of meteorological extremes over the southeast of the Iberian Peninsula. *Remote Sens.* **2018**, *10*, 1023. [CrossRef]
5. Cersosimo, A.; Larosa, S.; Romano, F.; Cimini, D.; Di Paola, F.; Gallucci, D.; Gentile, S.; Geraldi, E.; Teodosio Nilo, S.; Ricciardelli, E.; Ripepi, E.; Viggiano, M. Downscaling of satellite OPEMW surface rain intensity data. *Remote Sens.* **2018**, *10*, 1763. [CrossRef]
6. 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]

7. 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]
8. 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]
9. Ivanov, S.; Michaelides, S.; Ruban, I. Mesoscale resolution radar data assimilation experiments with the Harmonie model. *Remote Sens.* **2018**, *10*, 1453. [CrossRef]
10. Khan, S.; Maggioni, V. Assessment of level-3 gridded Global Precipitation Mission (GPM) products over oceans. *Remote Sens.* **2019**, *11*, 255. [CrossRef]
11. 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]
12. Le, H.M.; Sutton, J.R.P.; Du Bui, D.; Bolten, J.D.; Lakshmi, V. Comparison and bias correction of TMPA precipitation products over the lower part of Red-Thai Binh River Basin of Vietnam. *Remote Sens.* **2018**, *10*, 1582. [CrossRef]
13. Lu, D.; Yong, B. Evaluation and hydrological utility of the latest GPM IMERG V5 and GSMaP V7 precipitation pProducts over the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 2022. [CrossRef]
14. Martins Costa do Amaral, L.; Barbieri, S.; Vila, D.; Puca, S.; Vulpiani, G.; Panegrossi, G.; Biscaro, T.; Sanò, P.; Petracca, M.; Marra, A.; Gosset, M.; Dietrich, S. Assessment of ground-reference data and validation of the H-SAF precipitation products in Brazil. *Remote Sens.* **2018**, *10*, 1743. [CrossRef]
15. Michot, V.; Vila, D.; Arvor, D.; Corpetti, T.; Ronchail, J.; Funatsu, B.; Dubreuil, V. Performance of TRMM TMPA 3B42 V7 in replicating daily rainfall and regional rainfall regimes in the Amazon basin (1998–2013). *Remote Sens.* **2018**, *10*, 1879. [CrossRef]
16. Omranian, E.; Sharif, H.O.; Tavakoly, A.A. How well can Global Precipitation Measurement (GPM) capture hurricanes? Case study: Hurricane harvey. *Remote Sens.* **2018**, *10*, 1150. [CrossRef]
17. Rahman, K.; 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]
18. Ramsauer, T.; Weiß, T.; Marzahn, P. Comparison of the GPM IMERG final precipitation product to RADOLAN weather radar data over the topographically and climatically diverse Germany. *Remote Sens.* **2018**, *10*, 2029. [CrossRef]
19. Rehman, A.; Chishtie, F.; Qazi, W.; Ghuffar, S.; Fatima, I. Evaluation of three-hourly TMPA rainfall products using telemetric rain gauge observations at Lai Nullah basin in Islamabad, Pakistan. *Remote Sens.* **2018**, *10*, 2040. [CrossRef]
20. Retalis, A.; Katsanos, D.; Tymvios, F.; Michaelides, S. Validation of the first years of GPM operation over Cyprus. *Remote Sens.* **2018**, *10*, 1520. [CrossRef]
21. Ricciardelli, E.; Di Paola, F.; Gentile, S.; Cersosimo, A.; Cimini, D.; Gallucci, D.; Geraldi, E.; Larosa, S.; Nilo, S.; Ripepi, E.; Romano, F.; Viggiano, M. Analysis of Livorno heavy rainfall event: examples of satellite-based observation techniques in support of numerical weather prediction. *Remote Sens.* **2018**, *10*, 1549. [CrossRef]
22. Rozante, J.R.; Vila, D.A.; Chiquetto, J.B.; De A. Fernandes, A.; Alvim, D.S. Evaluation of TRMM/GPM blended daily products over Brazil. *Remote Sens.* **2018**, *10*, 882. [CrossRef]
23. Rysman, J.F.; Panegrossi, G.; Sanò, P.; Marra, A.C.; Dietrich, S.; Milani, L.; Kulie, M.S. SLALOM: An all-surface snow water path retrieval algorithm for the GPM microwave imager. *Remote Sens.* **2018**, *10*, 1278. [CrossRef]
24. Sanò, P.; Panegrossi, G.; Casella, D.; Marra, A.C.; D'Adderio, L.P.; Rysman, J.F.; Dietrich, S. The passive microwave neural network precipitation retrieval (PNPR) algorithm for the CONICAL scanning Global Microwave Imager (GMI) radiometer. *Remote Sens.* **2018**, *10*, 1122. [CrossRef]
25. 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]
26. Senent-Aparicio, J.; López-Ballesteros, A.; Pérez-Sánchez, J.; Segura-Méndez, F.J.; Pulido-Velazquez, D. Using multiple monthly water balance models to evaluate gridded precipitation products over peninsular Spain. *Remote Sens.* **2018**, *10*, 922. [CrossRef]
27. 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]

28. Tapiador, F.J.; Marcos, C.; Navarro, A.; Jiménez-Alcázar, A.; Galdón, R.M.; Sanz, J. Decorrelation of satellite precipitation estimates in space and time. *Remote Sens.* **2018**, *10*, 752. [CrossRef]
29. Tapiador, F.J.; Moreno, R.; Navarro, A.; Jiménez, A.; Arias, E.; Cazorla, D. Variability of microwave scattering in a stochastic ensemble of measured rain drops. *Remote Sens.* **2018**, *10*, 960. [CrossRef]
30. Tapiador, F.J.; Navarro, A.; Marcos, C.; Moreno, R. Estimates of the change in the oceanic precipitation off the coast of Europe due to increasing greenhouse gas emissions. *Remote Sens.* **2018**, *10*, 1198. [CrossRef]
31. Tsarpalis, K.; Papadopoulos, A.; Mihalopoulos, N.; Spyrou, C.; Michaelides, S.; Katsafados, P. The implementation of a mineral dust wet deposition scheme in the GOCART-AFWA module of the WRF model. *Remote Sens.* **2018**, *10*, 1595. [CrossRef]
32. Varlas, G.; Anagnostou, M.N.; Spyrou, C.; Papadopoulos, A.; Kalogiros, J.; Mentzafou, A.; Michaelides, S.; Baltas, E.; Karymbalis, E.; Katsafados, P. A multi-platform hydrometeorological analysis of the flash flood event of 15 November 2017 in Attica, Greece. *Remote Sens.* **2018**, *11*, 45. [CrossRef]
33. 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]
34. Wu, Z.; Xu, Z.; Wang, F.; He, H.; Zhou, J.; Wu, X.; Liu, Z. Hydrologic evaluation of multi-source satellite precipitation products for the upper Huaihe River Basin, China. *Remote Sens.* **2018**, *10*, 840. [CrossRef]
35. Xie, Y.; Shi, J.; Fan, S.; Chen, M.; Dou, Y.; Ji, D. Impact of radiance data assimilation on the prediction of heavy rainfall in RMAPS: A case study. *Remote Sens.* **2018**, *10*, 1380. [CrossRef]
36. Zhang, C.; Chen, X.; Shao, H.; Chen, S.; Liu, T.; Chen, C.; Ding, Q.; Du, H. Evaluation and intercomparison of high-resolution satellite precipitation estimates-GPM, TRMM, and CMORPH in the Tianshan Mountain Area. *Remote Sens.* **2018**, *10*, 1543. [CrossRef]
37. Zhang, R.; Wang, Z.; Hilburn, K. Tropical cyclone rainfall estimates from FY-3B MWRI brightness temperatures using the WS algorithm. *Remote Sens.* **2018**, *10*, 1770. [CrossRef]
38. Zhang, X.; Anagnostou, E.N.; Schwartz, C.S. NWP-based adjustment of IMERG precipitation for flood-inducing complex terrain storms: Evaluation over CONUS. *Remote Sens.* **2018**, *10*, 642. [CrossRef]
39. Zhang, Y.; Li, Y.; Ji, X.; Luo, X.; Li, X. Evaluation and hydrologic validation of three satellite-based precipitation products in the upper catchment of the Red River basin, China. *Remote Sens.* **2018**, *10*, 1881. [CrossRef]

remote sensing

MDPI

Article

Advancing Precipitation Estimation and Streamflow Simulations in Complex Terrain with X-Band Dual-Polarization Radar Observations

Marios N. Anagnostou [1,2,3,*], Efthymios I. Nikolopoulos [4], John Kalogiros [1], Emmanouil N. Anagnostou [4], Francesco Marra [5], Elisabeth Mair [6], Giacomo Bertoldi [6], Ulrike Tappeiner [6,7] and Marco Borga [8]

1 National Observatory of Athens, IERSD, 15236 Athens, Greece; jkalog@noa.gr
2 Department of Environmental Engineering, School of Engineering, DUTH, 67100 Xanthi, Greece
3 Department of Water Resources, School of Civil Engineering, NTUA, 10682 Athens, Greece
4 Civil and Environmental Engineering, School of Engineering, University of Connecticut, Hartford, CT 06269, USA; enikolop@engr.uconn.edu (E.I.N.); manos@engr.uconn.edu (E.N.A.)
5 Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem 9190401, Israel; marra.francesco@mail.huji.ac.il
6 Institute for Alpine Environment, Eurac research, 39100 Bolzano, Italy; mair_eli@yahoo.de (E.M.); giacomo.bertoldi@eurac.edu (G.B.); ulrike.tappeiner@eurac.edu or ulrike.tappeiner@uibk.ac.at (U.T.)
7 Department of Ecology, University of Innsbruck, Innsbruck A-6020, Austria
8 Department of Land, Environment, Agriculture and Forestry, University of Padova, 35122 Padova, Italy; marco.borga@unipd.it
* Correspondence: managn@noa.gr or sifneos@live.com; Tel.: +30-697-11-32-59

Received: 26 June 2018; Accepted: 6 August 2018; Published: 10 August 2018

Abstract: In mountain basins, the use of long-range operational weather radars is often associated with poor quantitative precipitation estimation due to a number of challenges posed by the complexity of terrain. As a result, the applicability of radar-based precipitation estimates for hydrological studies is often limited over areas that are in close proximity to the radar. This study evaluates the advantages of using X-band polarimetric (XPOL) radar as a means to fill the coverage gaps and improve complex terrain precipitation estimation and associated hydrological applications based on a field experiment conducted in an area of Northeast Italian Alps characterized by large elevation differences. The corresponding rainfall estimates from two operational C-band weather radar observations are compared to the XPOL rainfall estimates for a near-range (10–35 km) mountainous basin (64 km^2). In situ rainfall observations from a dense rain gauge network and two disdrometers (a 2D-video and a Parsivel) are used for ground validation of the radar-rainfall estimates. Ten storm events over a period of two years are used to explore the differences between the locally deployed XPOL vs. longer-range operational radar-rainfall error statistics. Hourly aggregate rainfall estimates by XPOL, corrected for rain-path attenuation and vertical reflectivity profile, exhibited correlations between 0.70 and 0.99 against reference rainfall data and 21% mean relative error for rainfall rates above 0.2 mm h^{-1}. The corresponding metrics from the operational radar-network rainfall products gave a strong underestimation (50–70%) and lower correlations (0.48–0.81). For the two highest flow-peak events, a hydrological model (Kinematic Local Excess Model) was forced with the different radar-rainfall estimations and in situ rain gauge precipitation data at hourly resolution, exhibiting close agreement between the XPOL and gauge-based driven runoff simulations, while the simulations obtained by the operational radar rainfall products resulted in a greatly underestimated runoff response.

Keywords: X-band radar; dual-polarization; precipitation; complex terrain; runoff simulations

1. Introduction

Natural disasters occurring in mountainous areas worldwide are most often associated with hazards like flash floods, debris flows, and landslides, which are responsible for fatalities, property losses, and environmental degradation [1,2]. The triggering process of those hazards depends on several factors including hydrological, geomorphological, and geotechnical features [3], but indisputably, the primary triggering factor is heavy rainfall interacting with a steep topography. Given the prospect of increase in frequency and intensity of heavy precipitation events in the context of climate change [4,5], the risk associated with rainfall-induced hazards is expected to increase in the future.

Early warning systems are essential elements in the development of disaster preparedness and response strategies [6,7]. The effectiveness of early warning procedures depends highly on the accuracies of the precipitation monitoring systems involved [4,8,9]. The impact of accurate representation of rainfall spatial variability on simulating the flash flood response has been demonstrated by several studies [10–12]. Moreover, recent studies [13,14] have shown the importance of accurate rainfall estimates at small spatial scales for the prediction of debris flow/landslide initiation. Therefore, to improve the accuracy of flood hazard warning systems, it is critical to acquire accurate estimates of precipitation rates at a high resolution.

The quantification of precipitation over mountainous areas, where these phenomena typically occur, is rather complex, and topography plays a dominant role in shaping precipitation variability [15]. Current operational rainfall monitoring systems from national weather radar networks do not provide sufficiently accurate measurements of precipitation variability over mountainous areas [16]. Oftentimes, storms developing in mountainous regions are affected by low-level orographic enhancement of precipitation that affects both the small-scale variability and intensity of rain rates and creates a great challenge to Quantitative Precipitation Estimation (QPE) by weather radar observations [17,18]. For operational radar networks, in particular, the monitoring of precipitation at long ranges is associated with precipitation detection issues including partial, or total, beam blockage of the lower beam elevations or overshooting of low-level convection signatures by the upper elevation sweeps, which leads to significant range dependent errors in the quantification of surface rainfall [19,20]. In addition, melting snow in widespread storm systems that typically result in high accumulations from persistent stratiform precipitation may substantially increase the threat of localized flash floods in mountainous basins [21].

Past studies [22–29] have shown that locally deployed dual-polarization X-band radar systems can contribute higher resolution rain rate estimations and improved rainfall quantification accuracies than the lower frequency (C-band and S-band) long-range operational radar systems. These short-range radar systems could potentially be used to fill in coverage gaps of operational weather radar networks, which is particularly important for advancing early warning of precipitation driven hydrological hazards (flash floods, landslides, debris flows, etc.) in urban and small mountainous basins [24,25,30]. A key drawback of X-band radar wavelength in rainfall estimation is the severe rain-path attenuation, but dual-polarization radar technology can now mitigate this effect through the development of robust attenuation correction schemes [31–41]. As a result, progress in the X-band dual-polarization rainfall estimation has enhanced the applicability of such radar (typically mobile) systems for flood modeling applications in urban areas [42–45] and in complex terrain basins [22–25,31,46,47].

The study's objective is to evaluate the QPE obtained from locally deployed X-band dual-polarization (XPOL) observations and associated accuracies in hydrologic simulations of small-scale mountainous basins compared to precipitation estimates derived from near-range operational radar (C-band) observations [48]. This study is based on a field experiment conducted in the framework of HyMeX (Hydrological cycle in a Mediterranean experiment) Special Observation Period One (SOP 1) in the summer-fall period of 2012 and 2014 over a mountainous region (above 2000 m elevation ranges) in the Eastern Italian Alps. Quantitative comparisons of precipitation estimate and associated hydrologic simulations between the locally deployed XPOL from this field experiment

and near-range operational C-band radars are used to highlight the benefits of locally deployed short-range X-band dual-polarization radar observations in terms of small-scale rainfall and runoff estimations. A description of the study area, the data collected and the setup of the field experiment are provided in the next section. The XPOL rainfall estimation algorithm and evaluation results are discussed in Section 2. Comparison of error statistics between XPOL and the operational C-band rainfall products and corresponding errors in hydrologic simulations are provided in Sections 3.1–3.3, respectively. Finally, the main conclusions are summarized in Section 4.

2. Instruments and Methodology

In the framework of HyMeX Special Observation experiment, the National Observatory of Athens (NOA) XPOL mobile radar was deployed in the northwestern part of Upper Adige river basin, in Eastern Italian Alps in July through October periods of 2012 and 2014 (Figure 1).

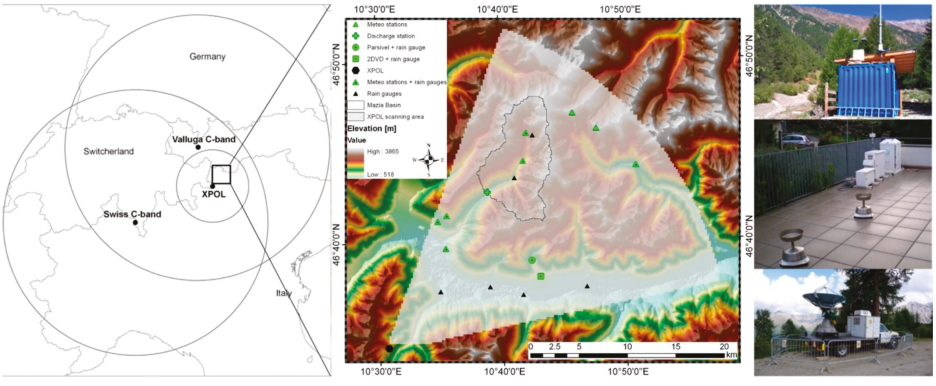


Figure 1. The topography of the experimental area showing the coverage of the XPOL radar encompassing the experimental basin, the various in situ instruments, and the coverage of the two operational network radars. Pictures from the deployed sensors are also shown (at the lower right is the XPOL radar, in the middle are two NOA gauges and the 2D-video disdrometer and at the top are the Parsivel disdrometer and Micro-Rain-Radar of NASA).

XPOL was deployed at an elevation of 2060 m (46°34′10.56″, 10°30′41.7594″) near range (40 km) of an experimental mountainous basin exhibiting dense in situ observations [49–51], which is part of the Long Term (Socio) Ecological Research site Matsch/Mazia (http://lter.eurac.edu/en). The upper Saldur catchment is located, with reference to its outlet, 10 km northeast from the XPOL site (Figure 1) and comprises the study area for the hydrologic analysis presented in this work. The basin covers a highly complex terrain area of 62 km^2 with elevations ranging from about 1400 to 3700 m a.s.l. and an average elevation of about 2500 m a.s.l. At the upper part of the basin, a significant portion of the area (2.8 km^2 in 2013) [52] is covered by glaciers. Due to its inner-alpine position, the mean annual precipitation in the lower valley is lower than 530 mm at 1600 m a.s.l. (Source: Mazia meteorological and discharge measurement stations of the Hydrographic Office South Tyrol). However, a strong lapse rate leads to more than 1000 mm year^{-1} at the higher altitudes of the basin, which touch the Alps main divide [53]. Precipitation typically occurs as snowfall from November to late April, but snowstorms can also occur during the summer and early fall at the higher elevations. During the late summer-early fall (the period of this experiment) precipitation typically originates from convective rainfall [54].

During the experimental periods, rainfall data were collected from the XPOL and a dense network of in situ sensors (rain gauges, meteorological stations, and disdrometers) contributed by the NOA, Eurac research (Bozen/Bolzano), and the hydrologic office of the Autonomous Province of Bolzano. Available instrumentation used in this study, shown in Figure 1, includes one 2D-video disdrometer

(2DVD) available only in 2012 period, one laser type Parsivel disdrometer (APU), fourteen rain gauges, and seven meteorological stations with the temporal resolution of 15 min [55]. Three of the rain gauge stations also provided snow depth data. Since most of the rain gauges are in high altitudes and none of them are heated, liquid precipitation is distinguished from snow using air temperature and snow-depth observations in order to avoid rainfall underestimations [56].

The XPOL radar was operated during rain events first in plan position indicator (PPI) mode taking measurements in a sector scan of 60°, at 0.5°, 1.5°, 2.5°, 3.5°, and 5° elevation sweeps with a range resolution of 120 m for a total range of 35 km. The radar also operated in a range height indicator (RHI) mode taking RHI measurements over the two disdrometer sites. The antenna rotation rate was 6 deg s^{-1} for PPI and 3 deg s^{-1} for RHI mode. The time period for a full volume scan including the RHIs was less than 3 min.

The study domain is also covered by two operational C-band radar networks, one from Austria (~60 km from study basin) and a second from Switzerland (~150 km from study basin), providing surface QPE maps. The Austrian Meteo Service network covers the study area with a C-band weather radar located at the Valluga Mountain (2809 m a.s.l.); hereafter named "Valluga". The radar covers the western parts of the Austrian Alps in the borders with the Italian Alps (shown in Figure 1). The lowest usable elevation is −0.5°, and a volume scan consists of 16 elevation sweeps with a duration of 5 min [57]. Technical details of the Valluga C-band radar are listed in Table 1. The rainfall maps from this radar are based on an algorithm described in [58,59]. Surface QPE maps are available at 1 km grid resolution and hourly accumulations.

Table 1. The technical specifications of the three C-band radars from for MeteoSwiss network and the Valluga radar of the Austrian weather radar operational network.

	1. Swiss Radars	2. Valluga Radar
Operating Frequency	C-band (5.4 GHz)	C-band (5.625 GHz)
Pulse Length	0.5 μs	0.8 μs
PRF	variable from 600 to 1500 Hz	1000 Hz
Polarization	Dual	Dual
Dynamic Range	105 dB	108 dB
Max Range	246 km	120
Angular Resolution	1° with 20 elevations per 5 min	1° with 14 elevations per 5 min
Range Resolution	83 m	125 m
Beamwidth	~1°	~0.92°

The second operational radar rainfall product over the study area is based on estimates from the MeteoSwiss weather radar network covering the area [60]; specifically, the MeteoSwiss radar observations over the study area are based on three dual-polarization C-band radars located at Mt Lema, La Dôle, and Albis in Switzerland (technical characteristics described in Table 1). The rainfall product from the MeteoSwiss operational radar network (hereafter named "Swiss") is based on an operational rainfall estimation algorithm called "CombiPrecip" [61–63] that provides hourly accumulated gauge-adjusted/radar merging rainfall maps at a 1 km grid resolution.

Table 2 summarizes the precipitation events selected in this study during the two experimental periods and the corresponding availability of sensor observations. The table also reports the average rainfall accumulation (in mm) reported from rain gauges deployed in the Saldur catchment. Ten out of the fifteen recorded events over the two experimental periods constitute the dataset of this study. The criteria for selecting these events are the availability of coincident operational C-band radar and XPOL rainfall datasets, in situ observations, and event total rainfall accumulations exceeding 10 mm. The "Swiss" radar rainfall product was used in 2014 and the 25 August 2012 event, while, for the rest of the selected storm events, we used rainfall products from the Valluga radar (Austrian radar network).

Although no major flooding occurred during the two field experiments in the study area, due to the moderate precipitation intensity of the storm events, the streamflow observations (made available

from the Hydrologic Office of the Autonomous Province of Bolzano) show that two of the recorded events were associated with basin response adequate to support flood modeling. Therefore, in addition to the precipitation data, discharge observations for three different locations at the Saldur basin are used in this study to verify the model applied for hydrologic simulations.

Table 2. The selected rain events for the two observations periods in 2012 and 2014. In this table, "Total average—gauge rain indicates the surface basin total (for the time period) average rainfall".

Time Period (UTC) (day/month/year)	Total Average–Gauge Rain (mm)	Available Observations
06/08/12 17:00–06/08/12 21:00	10	Rain gauges (all sites), 2DVD, Parsivel, XPOL, Valluga
25/08/12 16:00–25/08/12 24:00	21	Rain gauges (all sites), Parsivel, XPOL, Swiss Radar Network
10/09/12 08:00–12/09/12 21:00	16	Rain gauges (all sites), 2DVD, Parsivel, MRR, XPOL, Valluga
24/09/12 07:00–27/09/12 21:30	28	Rain gauges (all sites), 2DVD, Parsivel, MRR, XPOL, Valluga
29/09/12 07:00–30/09/12 06:30	11	Rain gauges (all sites), 2DVD, Parsivel, MRR, XPOL, Valluga
04/08/14 15:00–05/08/14 04:00	11	Rain gauges (all sites), Parsivel, XPOL, Swiss Radar Network
09/08/14 08:00–09/08/14 15:00	14	Rain gauges (all sites), Parsivel, XPOL, Swiss Radar Network
12/08/14 06:00–13/08/14 22:00	37	Rain gauges (all sites), Parsivel, XPOL, Swiss Radar Network
29/08/14 20:00–01/09/14 08:00	10	Rain gauges (all sites), Parsivel, XPOL, Swiss Radar Network
15/09/14 14:00–16/09/14 23:00	13	Rain gauges (all sites), Parsivel, XPOL, Swiss Radar Network

XPOL Radar Rainfall Estimation Algorithm

This section describes the microphysical algorithm (named SCOP-ME) of Reference [64] applied on the XPOL radar observations. XPOL reflectivity observations were corrected for partial beam blockage and measurement biases prior to applying the precipitation retrieval algorithm. For beam blockage estimation, high-resolution terrain information and a three-dimensional model of the radar beam was applied. This information is used to exclude from further processing highly occluded areas or to correct reflectivity from areas with a minor (<50%) occlusion.

Next, the calibration of the XPOL radar echo signal (horizontal and differential reflectivity) was determined for the correct application of the XPOL radar algorithms. This calibration is assessed by comparing XPOL measurements to reference radar parameters (reflectivity and differential reflectivity) determined using scattering (T-matrix) routines applied on the drop size distribution (DSD) measurements from the 2DVD and Parsivel disdrometers deployed in the study area (both at ~17 km range and directly visible from XPOL). The lowest antenna elevation (0.5°) was used because XPOL observations at this elevation are closest in altitude to the disdrometers, while there is no significant beam blockage or ground clutter affecting the XPOL measurements. The mean bias determined from this process was subsequently removed from the radar data, as is described in more detail in the Kalogiros et al. [64], and it was found to be 3.5 dB and −0.2 dB for the horizontal and differential reflectivity, respectively. These values are close to the calibration biases reported from previous XPOL deployments [22,37].

Figure 2a,b shows scatter plots of the rain-path attenuation corrected and bias-adjusted of XPOL horizontal (Z_{hR}) and differential (Z_{drR}) reflectivity versus the corresponding Z_{hD} and Z_{drD} parameters determined by rain spectra observations from the two disdrometers (2DVD—black dot and Parsivel—gray dot) in 2012 (Figure 2a) and only the Parsivel in 2014 (Figure 2b). The bias calibration of Z_{hR} and Z_{drR} are made using long-term disdrometer data as described in Kalogiros et al. [64] and Anagnostou et al. [37]. The calibration of Z_{drR} is also improved using a real-time method which is based on average Z_h-Z_{dr} relations. It is actually preformed in every radar scan using a Z_{dr} vs. Z_h average relation which is dynamically corrected for each scan, as described in more details in Reference [64]. The correlation coefficient of XPOL radar parameters compared to both disdrometers (2DVD and Parsivel) and for both observation years are 0.80 and 0.65 for the horizontal and differential reflectivity, respectively, and the mean relative error (after removing the corresponding biases) is less than 3% for both reflectivity types. Even though the scatter plots (especially the Z_{dr} plot) look noisy due factors like measurements errors, the difference in height between the radar volume and the disdrometer, and the difference in spatial resolution of the two systems (radar volume versus

point measurement), what is important is that the mean bias can be reliably estimated despite the scattering. Similar scatter has been observed in past studies over complex terrain [22,23], In addition, as it was mentioned above, the Z_{dr} calibration is actually made in every radar scan using a Z_{dr} vs. Z_h average relation [64]. Next, we briefly describe the rainfall retrieval algorithm applied to bias-adjusted XPOL parameters.

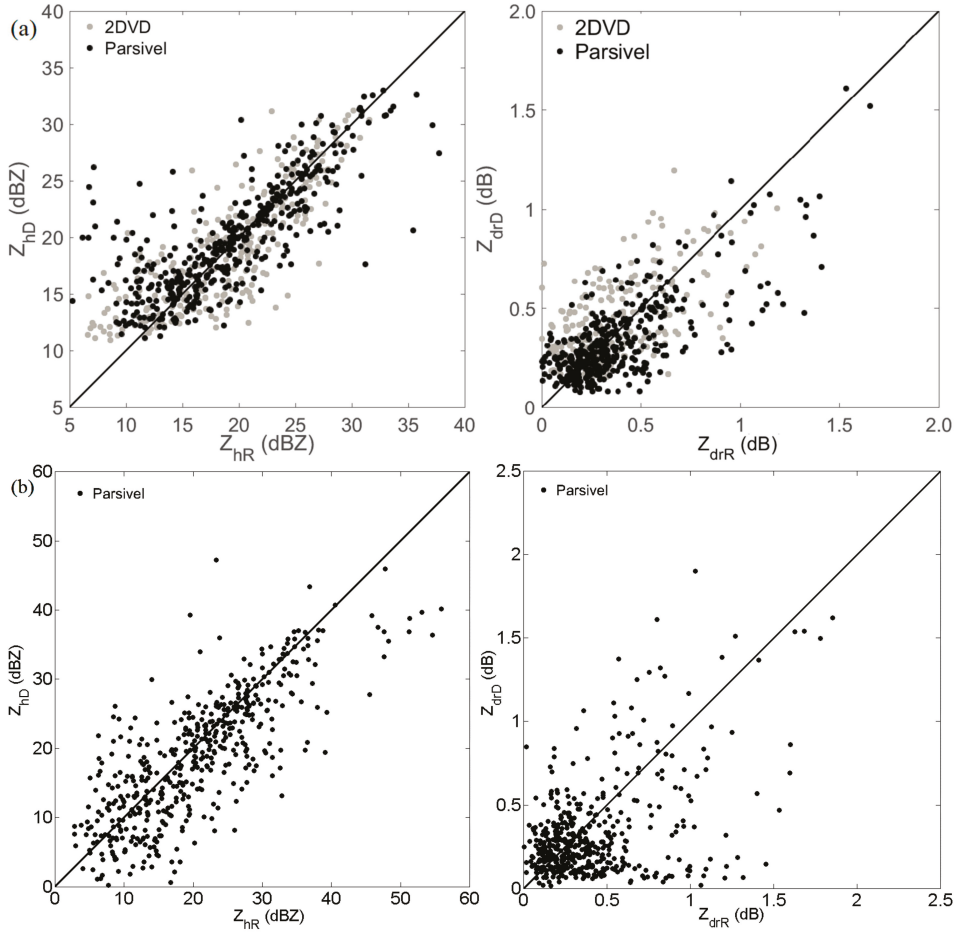


Figure 2. The scatter plots between the calibrated radar reflectivity Z_{hR} and differential reflectivity Z_{drR} at the position of the disdrometers (0.5° elevation angle of the radar antenna) and the corresponding reflectivities (Z_{hD} and Z_{drD}) estimated from the disdrometers for (**a**) 2012 and (**b**) 2014 observations.

The first step of the algorithm is the identification of the rain precipitation regime in each radar ray and application of the SCOP algorithm [36,64]. The path attenuation of a radar signal is significant, especially for high-frequency radars (like X-band). For the correction of path attenuation in rain, the SCOP algorithm is used. This algorithm is a self-consistent polarimetric algorithm, based on the parameterizations of the specific attenuation coefficients and backscattering phase shift in the rain, derived by Kalogiros et al. [64], and applied with an iterative scheme to separate radar rays [64]. Comparison with in situ disdrometer and rain gauge data in past studies [35,36,64] has shown that

the attenuation correction procedure described in Reference [64] is more efficient than previous polarimetric attenuation correction algorithms [65,66].

Due to the high elevation of the XPOL deployment in this study, the radar beam frequently intercepted the snow-rain melting layer and the snow region of stratiform precipitation above the study area (40 km range). Thus, to properly estimate rainfall rate at ground level, a correction for the effect of the melting layer and the snow layer above it on the radar measurements was applied. The methodology followed here is an extension of Reference [67] technique applied to PPI scans using the apparent vertical profile of reflectivity (VPR) shape (hereafter called the apparent VPR correction) with temporally variable characteristics determined from the radar data itself [35]. The apparent VPR correction algorithm was used to identify the bright band zone and snow region above, and remove these effects, bringing the spatial variability of reflectivity closer to the actual values below the melting layer (i.e., in the rain layer near the ground provided that the melting layer has not reached the ground level).

After rain-path attenuation and VPR correction, we applied the rainfall estimation and microphysical algorithms defined by Equations (2)–(4d) in Reference [35]. To understand the effectiveness of the VPR correction in rainfall estimation, Figures 3 and 4 show the XPOL rainfall timeseries (before and after the VPR correction for the 1.5° and 3.5° elevation) compared to the corresponding timeseries of rainfall derived from the two disdrometers (i.e., Parsivel and 2DVD) for the rain event of 24 September 2012 and from the Parsivel disdrometer for the rain events of 9 and 12–13 August 2014. When the VPR correction reduces (increases) the rainfall values estimated by XPOL, the radar beam gate at the horizontal position of the in-situ sensor was located in the bright band (above the bright band and in the snow region). Despite the random differences shown in Figures 3b and 4b between the XPOL and in situ rainfall for the high elevation angle (3.5°), the comparison of accumulated rainfall depicted in Figures 3c and 4c, shows that the VPR correction significantly reduces the difference between radar and in situ ground observations. Removing this systematic error in accumulated rainfall is a critical factor in hydrological runoff modeling [68], which will be discussed later in the paper.

Figures 5 and 6 show the PPI of XPOL radar accumulated rainfall estimates without (upper two plots) and with VPR (lower two plots) correction from two high-elevation sweeps (2.5° and 3.5°). It is noted that the spatial structure of accumulated rainfall is similar between the two elevations after removing the bright band overestimation and snow underestimation based on the VPR correction procedure. In Figure 6, it can be seen that the bright band (melting layer) is at a longer range (and, thus, at higher height) than in Figure 5. In Figure 6 its base is at about 2800 m (radar height is about 2100 m), while in Figure 5 its base is at about 2400 m. More precisely, at the range where the base of the bright band is located in 2.5° elevation of Figure 6 the radar beam is completely blocked by the slope of the mountain (i.e., it is too low and interacts with the ground) in the center of the scanning sector and, thus, bright band detection and correction is very difficult for this elevation. The differences in rainfall field between the 2.5° and 3.5° elevations are found before the bright band range (i.e., below its height) and have nothing to do with the VPR correction because no VPR correction is applied below the bright band. These differences are probably due to some ground clutter in the 2.5° elevation PPI, which remained even after the application of the clutter filter. This is the reason that radar data from the 3.5° elevation was used, but not a higher elevation in order to limit the height of the radar volume above the surface and the effect of melting layer at close ranges, for further analysis (Section 3).

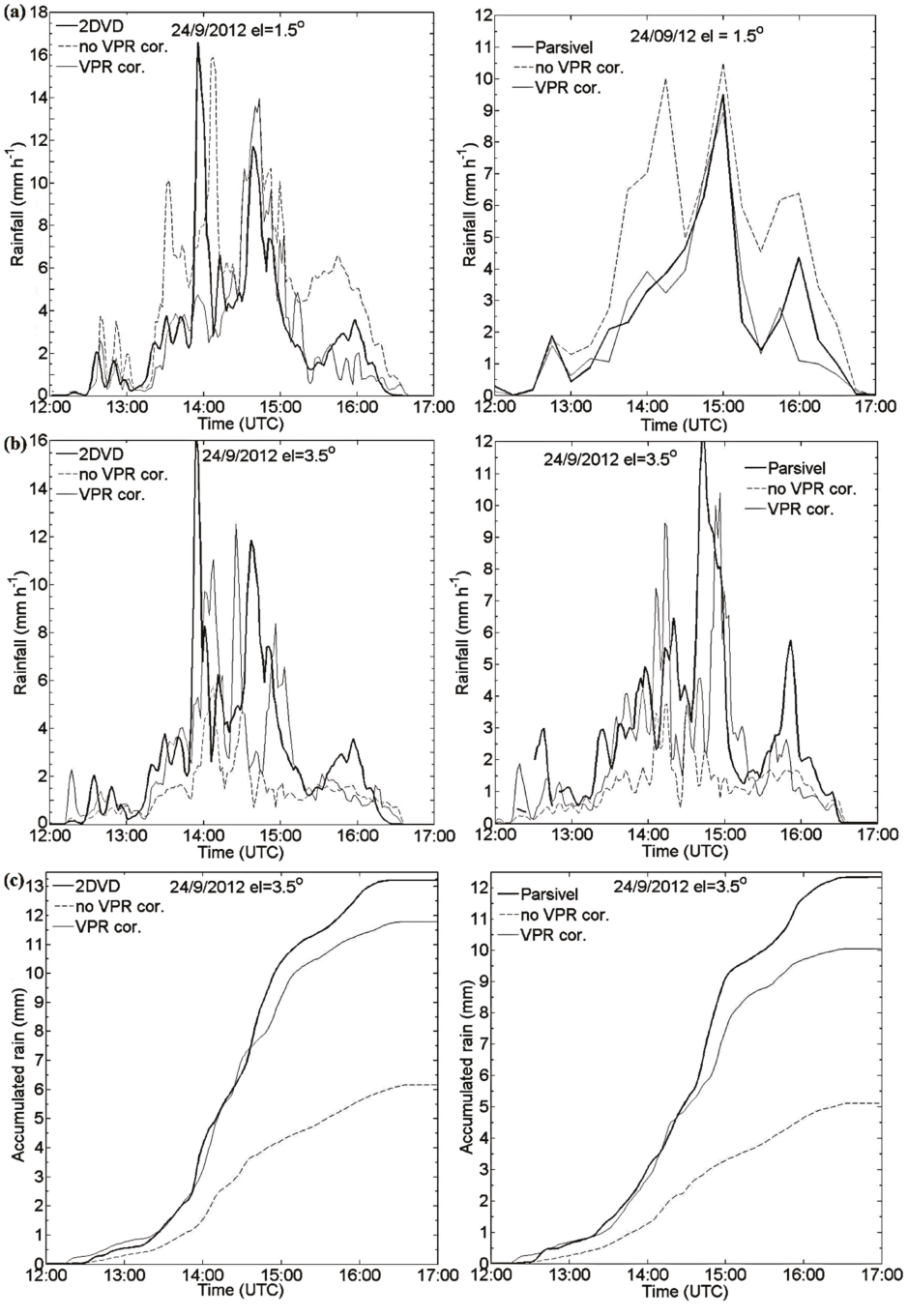


Figure 3. The application of the VPR correction algorithm to XPOL data at (**a**) 1.5° and (**b**,**c**) 3.5° elevation and comparison with in situ (2DVD and Parsivel) data on 24 September 2012.

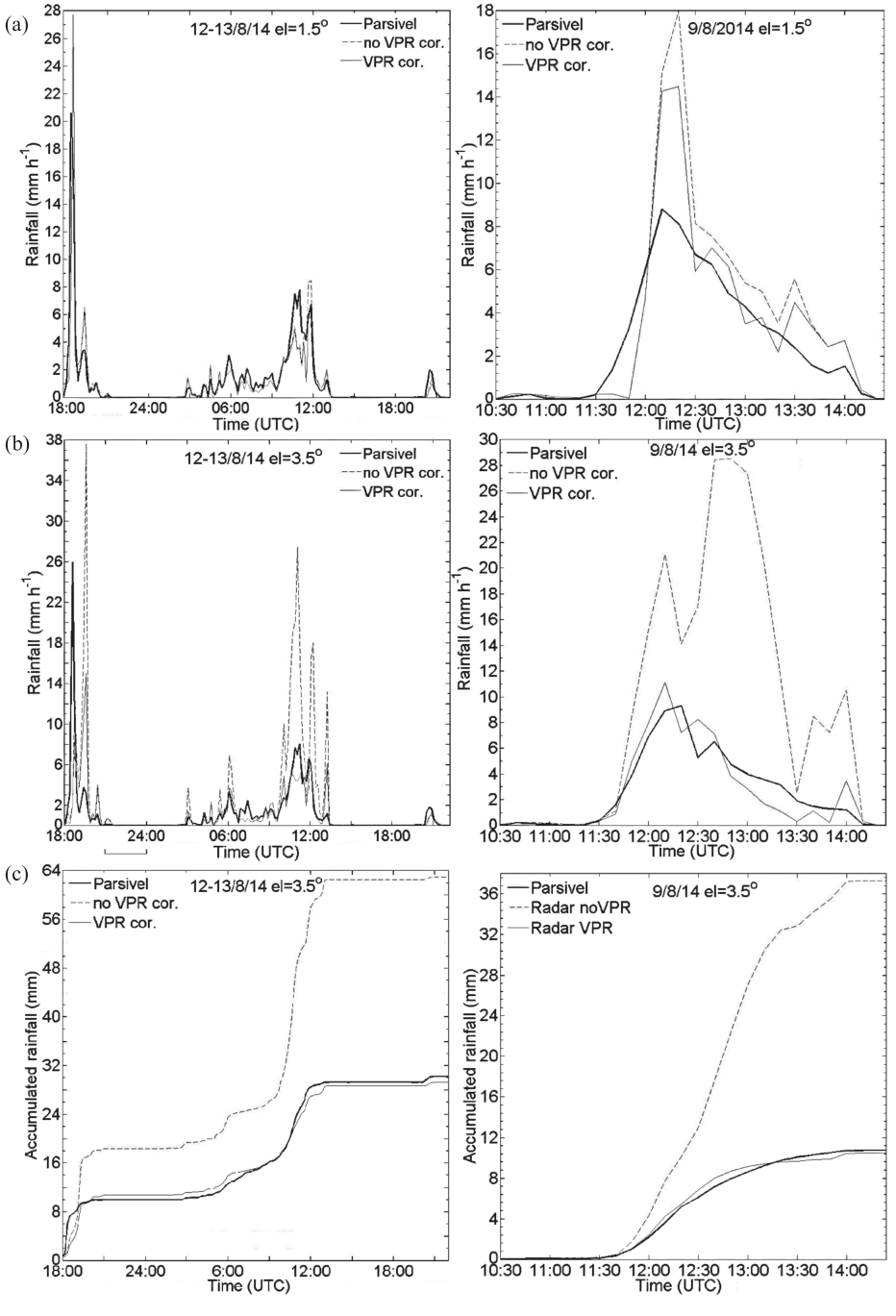


Figure 4. The same as Figure 3, but for the rain events of 12–13 and 9 August 2014 at (**a**) 1.5° and (**b**,**c**) 3.5°.

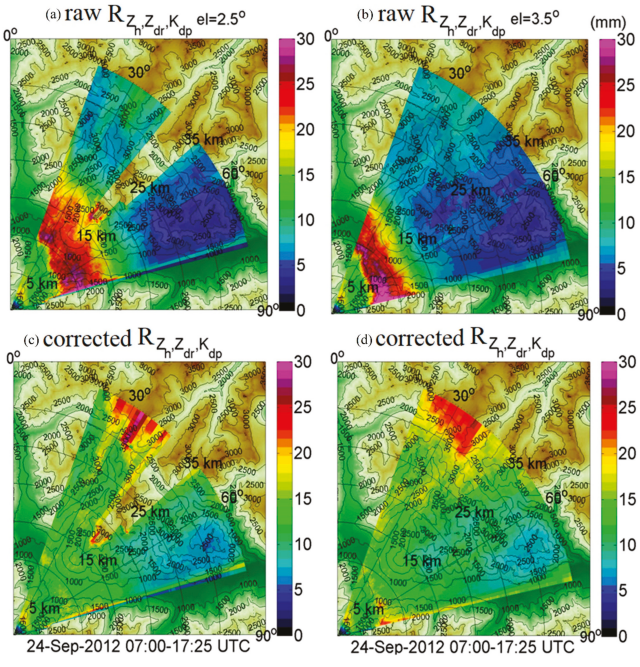


Figure 5. The PPI plots of accumulated rain estimated by the XPOL radar before (**a**,**b**) and after the VPR correction (**c**,**d**) at elevations 2.5° (**a**,**c**) and 3.5° (**b**,**d**) for the rain event of 24 September 2012.

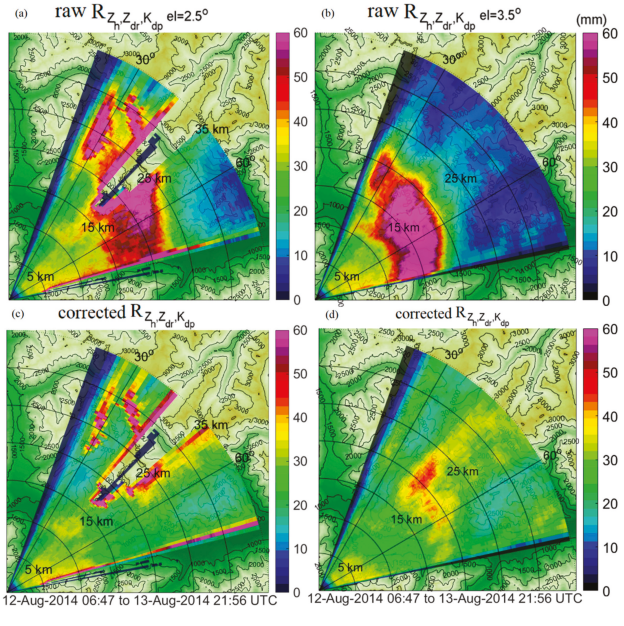


Figure 6. The same as Figure 5 but for the rain event of 12 August 2014, before (**a**,**b**) and after the VPR correction (**c**,**d**) at elevations 2.5° (**a**,**c**) and 3.5° (**b**,**d**).

Furthermore, scatterplots (Figure 7) of radar rainfall estimates before and after the VPR correction against disdrometer measurements for all the recorded rain events (even the rain events excluded from this study) for both field experiments indicate similar performance with a mean relative error change from about 50% overestimation before VPR correction to less than 10% after the application of the algorithm.

As a final step, XPOL radar rainfall estimates were converted to a Cartesian grid and aggregated to half-hourly rainfall accumulations (in mm h^{-1}). Linear interpolation was used to interpolate from the original polar grid to the Cartesian grid with a 200 × 200 m resolution. The gate resolution of XPOL was 120 m (as it is mentioned in the third paragraph of Section 2) and the angular resolution was 0.6°. Thus, the original polar grid resolution matches, on average, the Cartesian grid resolution in the area of the basins of interest. The XPOL rainfall estimation maps used in this work were processed at a high (200 × 200 m) and coarser (1 × 1 km) grid resolution, comparable with the "Austrian" and "Swiss" radar precipitation products. The 1 × 1 km Cartesian grid was made by averaging the 200 × 200 m grid cells contained in each cell of the coarser grid.

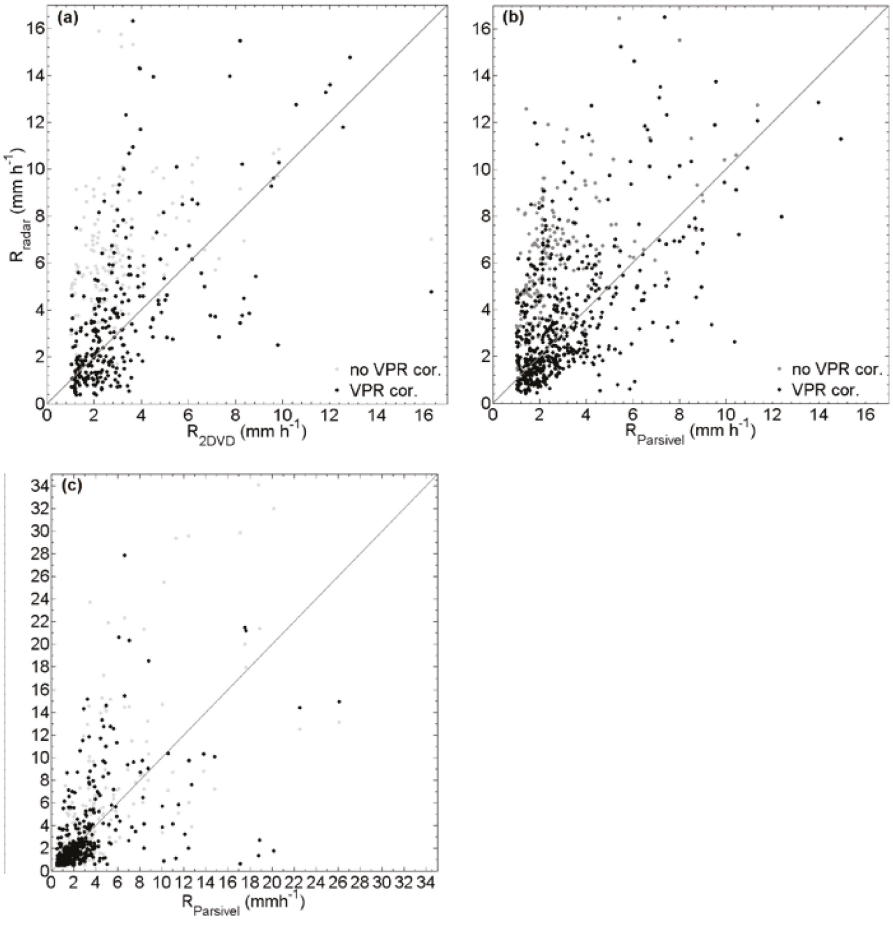


Figure 7. The scatter plots of rainfall rate, estimated from the radar (at the 1.5° elevation) without (light grey dots) and with VPR (black dots) correction against in situ measurements for the rain events of (**a**,**b**) 2012 and (**c**) 2014.

3. Results and Discussion

This section contrasts the quantitative evaluation of high-resolution XPOL rainfall estimates to the operational Swiss and Austrian radar network estimates using error statistics and scatter plots based on in situ rain gauge rainfall observations as a reference. In addition, this section shows a qualitative comparison between XPOL and the two operational radar rainfall products using storm total accumulation maps of two rainfall events (25 August 2012 and 12–13 August 2014) that triggered moderate peak runoff at the Saldur basin outlet.

3.1. Quantitative Evaluation of Radar-Rainfall Estimates

The statistical metrics for the evaluation of the radar-rainfall estimates include (1) the correlation coefficient between hourly radar-rainfall estimates and reference rainfall; (2) the unconditional bias ratio, which is defined as the ratio of storm event total radar rainfall estimates to the corresponding reference values; and (3) the conditional bias ratio, which is the same as in (2), but for values of the reference rainfall greater than 0.1 mm h^{-1}. The presented statistical analysis is evaluated based on hourly rainfall accumulation values (in mm h^{-1}) from a network of rain gauges divided into two groups, namely, gauges located inside the Saldur basin with a mean elevation at 2664 m (ranged from 1827 to 3500 m), hereafter named "high-elevation" gauges; and outside the Saldur basin with a mean elevation of 1095 m (ranged from 700 to 1490 m), hereafter named "low-elevation" gauges. The XPOL rainfall estimates are evaluated at both XPOL grid resolutions (200 m and 1 km) to show potential spatial aggregation effects on the error statistics, given that operational radar rainfall estimates are available at a 1-km grid resolution.

A summary of bulk statistical metrics (correlation coefficient, total unconditional and conditional—for reference rainfall > 0.1 mm h^{-1}—biases) of rainfall between estimates (i.e., the two C-band and XPOL radars) and reference (rain gauges) are presented in Table 3. We compared the performance metrics for both high and low elevation gauges. Analyzing the unconditional and conditional biases will provide a metric of rainfall detection error. According to these results XPOL rainfall estimates outperform the operational (C-band based) radar-rainfall products consistently for all events and gauge-elevation groups. Specifically, XPOL rainfall estimates exhibit correlations in the range of 0.70 to 0.99 for the low elevation and 0.62 to 0.99 for the high elevation gauge groups. Both Table 3 and Figure 8 results are for 1 h accumulated rainfall interpolated in an orthogonal grid of 200 m or 1 km resolution, respectively. Thus, a significant amount of spatial and temporal smoothing has been done, which means less random error and higher correlation with rain gauge data. It is important to note here that the 0.99 correlation from the 24/09 event is not statistically significant since the time duration was 5 h (see Figure 3) and, thus, only 5 points from this event were included in Table 3 and Figure 8, results.

Table 3. The statistical evaluation of the radar rainfall products; statistics in parenthesis represents the "low-elevation" reference in situ inside the Saldur basin.

	Correlation			Conditional Bias			Unconditional Bias		
Date/Radar (dd/mm/yy)	**C-Bands**	**XPOL-200 m**	**XPOL-1000 m**	**C-Bands**	**XPOL-200 m**	**XPOL-1000 m**	**C-Bands**	**XPOL-200 m**	**XPOL-1000 m**
06/08/12	0.65 (0.65)	0.82 (0.97)	0.83 (0.95)	0.69 (0.51)	0.95 (1.02)	0.93 (1.04)	0.37 (0.49)	1.26 (1.24)	1.22 (1.20)
25/08/12	0.69 (0.75)	0.89 (0.84)	0.89 (0.83)	0.32 (0,19)	1.11 (1.19)	1.12 (1.18)	0.27 (0.21)	1.08 (1.22)	1.10 (1.21)
10/09/12	0.71 (0.69)	0.68 (0.90)	0.70 (0.91)	0.35 (0.21)	0.77 (1.02)	0.77 (0.99)	0.35 (0.24)	0.81 (1.09)	0.82 (1.12)
24/09/12	0.68 (0.67)	0.94 (0.99)	0.91 (0.98)	0.34 (0.40)	1.06 (1.03)	1.09 (1.03)	0.24 (0.31)	1.08 (1.05)	1.11 (1.07)
29/09/12	0.43 (0.51)	0.82 (0.95)	0.81 (0.95)	0.91 (0.37)	0.98 (0.91)	1.01 (0.89)	0.34 (0.32)	1.16 (1.18)	1.18 (1.21)
04/08/14	0.53 (0.56)	0.94 (0.88)	0.93 (0.83)	1,96 (2.08)	0.71 (0.99)	0.67 (1.11)	1.96 (2.13)	0.76 (0.99)	0.71 (1.11)

Table 3. *Cont.*

	Correlation			Conditional Bias			Unconditional Bias		
Date/Radar (dd/mm/yy)	**C-Bands**	**XPOL-200 m**	**XPOL-1000 m**	**C-Bands**	**XPOL-200 m**	**XPOL-1000 m**	**C-Bands**	**XPOL-200 m**	**XPOL-1000 m**
09/08/14	0.69 (0.68)	0.83 (0.92)	0.77 (0.86)	0.72 (0.84)	0.74 (1.01)	0.68 (1.07)	0.75 (0.97)	0.74 (1.03)	0.68 (1.06)
12/08/14	0.68 (0.71)	0.81 (0.91)	0.80 (0.85)	0.58 (0.91)	1.17 (0.99)	1.31 (0.97)	0.58 (1.03)	1.21 (1.01)	1.35 (0.96)
29/08/14	0.52 (0.57)	0.62 (0.70)	0.54 (0.60)	0.70 (0.84)	1.69 (1.60)	2.06 (2.14)	0.68 (0.79)	2.01 (1.91)	2.40 (2.43)
15/09/14	0.66 (0.88)	0.83 (0.93)	0.77 (0.58)	2.38 (0.45)	2.52 (1.34)	2.81 (1.89)	2.50 (0.45)	2.71 (1.56)	3.10 (2.06)

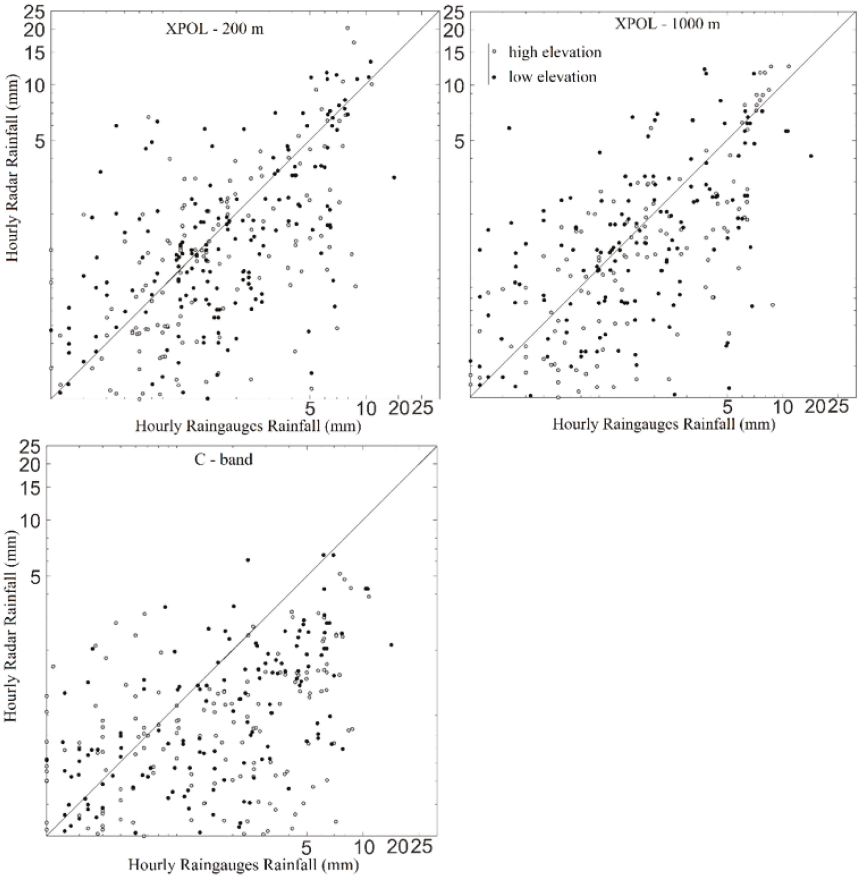


Figure 8. The scatter plots between the hourly rainfall estimation of XPOL and the large operational network radars versus the ground validation in situ sites separated into low and high elevation (inside the Saldur basin) gauge groups. The black line is the '1-1' relation.

On the other hand, the operational radar products show a significantly increased variation in performance with correlation values, being between 0.51 and 0.88 for the low elevation gauge group and 0.48 to 0.71 for the high elevation gauge group. The results for the conditional and unconditional bias ratios show that for the majority of cases, the majority of the operational radar rainfall products exhibit severe underestimation with the bias ratio (both conditional and unconditional) exhibiting

values ranging from 0.21 up to 0.91 for the low elevation gauge groups, and from 0.24 up to 0.97 for the high elevation gauge groups. However, there are some cases (i.e., 4 August 2014 and 15 September 2014 events) where the C-band radar rainfall overestimated the surface rainfall.

On the other hand, the XPOL estimates are associated with considerably reduced bias, compared to the C-band estimates, that, although it varies largely among the cases, is on average 30%. A distinct feature from the results reported in Table 3 is an apparent large overestimation of the XPOL rainfall estimates in the last two events of 2014 field experiment (29 August and 15 September) where it reached up to 3-times the rainfall from the reference. In these two events, although the XPOL rainfall estimates did not capture the peak rainfall rate values accurately they performed better for the lower rainfall rates. This can be explained due to the close ground melting layer and missed correction of the VPR algorithm to the rainfall estimate and to the "undercatch" problem of the ground reference (rain gauges) measure during low rainfall rates (<1 mm h^{-1}).

Figure 8 shows scatter plots of hourly rainfall (mm) estimates from the XPOL and the C-band radar products against the low and high elevation gauge-group hourly rainfall accumulations. It is noted that the XPOL rainfall estimates are almost unbiased in moderate to high rainfall values (i.e., >2 mm h^{-1}) for all cases. At the low-elevation gauge group, XPOL rainfall estimates exhibit a systematic overestimation, consistent with the bias statistics discussed above. Bias statistics presented in Table 3 and the scatter plots of Figure 8 suggest that a significant grid resolution effect on the error statistics is present. On the other hand, both Figure 8 and Table 3 results reveal that the longer-range operational C-band radar rainfall estimations systematically underestimated rainfall at both low and high elevation gauge groups. A point to note is that both of these C-band radar rainfall products include operational gauge-based adjustments, though located far away from our study area, while the XPOL rainfall polarimetric estimates are based on the disdrometer-driven radar parameter adjustments and a polarimetric technique that is independent of in situ rainfall measurements [64]. Nevertheless, these are the best operational radar rainfall products available for the study area (the Italian radar network is blind due to beam blockage) and represent in this study the baseline radar-rainfall performance characteristics.

3.2. Case Study Analysis of Radar Rainfall Spatial Pattern

In this section, the evaluation of the radar rainfall estimates is performed qualitatively by comparing the spatial patterns of rain accumulation for the two rain events identified previously as the ones that caused considerable runoff response in Saldur basin. The first event occurred on the 25th of August, 2012 and is classified as convective type. It lasted less than 10 h and resulted in rainfall accumulations greater than 25 mm. The second event occurred over a 35-h period on 12–13th August 2014 and was stratiform (widespread) precipitation, with an embedded convection in the northeast direction of the storm, associated with a higher event total rainfall accumulation (~35 mm) within the study domain (Figure 9). These events are used to demonstrate the differences in representation of spatial rainfall patterns between the operational (C-band) radar-rainfall products, the short-range/high-resolution XPOL measurements, and rain gauges. Figure 9a,b exhibits distinct differences in terms of the rainfall patterns captured by the two radar-rainfall products. Rainfall fields observed by the C-band radar networks are significantly underestimated and spatially smoothed relative to the precipitation fields retrieved by XPOL measurements. Observations from the visual interpretation of Figure 9 are further supported by the Taylor diagram [69] shown in Figure 10, which is a graphical representation of the statistical relationship between the point (pixel) radar-rainfall estimates (i.e., XPOL and the two C-band radars) and the respective one of the reference (i.e., rain gauge rainfall observations). The similarity between the two fields of the figure is quantified in terms of their correlation, their centered root-mean-square difference (cRMSD) and the amplitude of their variations (represented by their standard deviations).

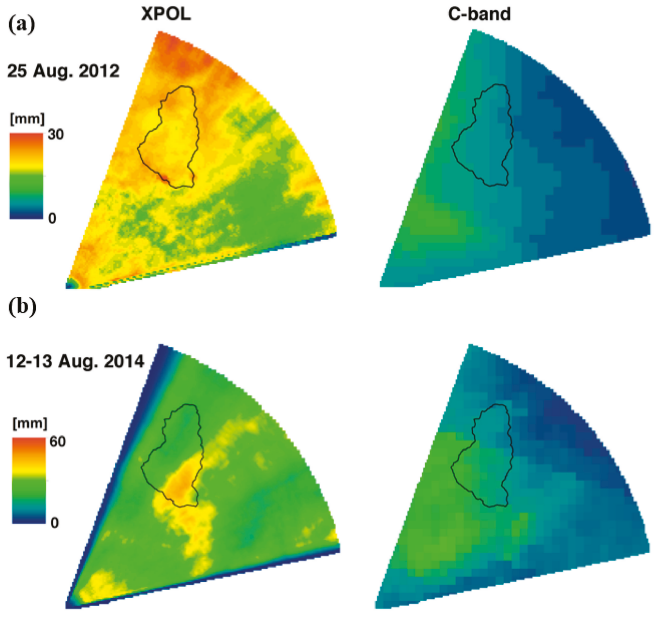


Figure 9. The storm total accumulation estimation maps (in mm) during the (**a**) 25 August 2012 and (**b**) the 12–13 August 2014 of Swiss network operation radar estimates (on the right panel) and XPOL estimates (on the left panels).

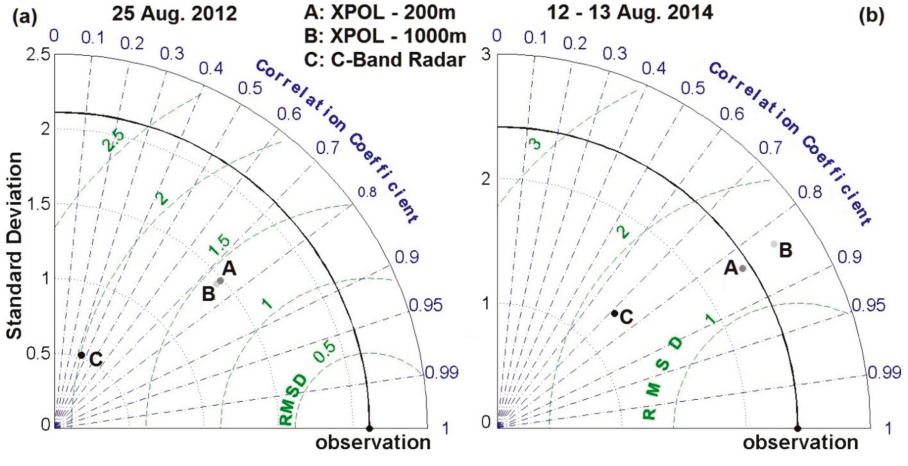


Figure 10. The Taylor diagram displaying a statistical comparison of three different daily accumulation radar precipitation estimates with in situ observations. The azimuthal angle represents correlation; the radial distance, the standard deviation; and the semicircles centered at the "observation" marker, the standard deviation of the errors. (**a**) 25 August 2012; (**b**) 12–13 August 2014.

Note that the means of the fields are subtracted out before computing their second-order statistics so the diagram does not provide information about overall biases, but solely characterizes the centered pattern error. The position of each letter appearing on the plot quantifies how closely the radar precipitation estimates match the reference. The cRMSD between the observed and the

reference patterns is proportional to the distance to the point on the x-axis identified as the reference "observation". For both events, we note a low correlation for the operational C-band estimates and a significant underestimation of the reference standard deviation. For instance, for the 25th of August (Figure 10a) event the correlation coefficient of the XPOL estimates (both 200 and 1000 m resolution) are nearly 0.75 compare to the C-band estimates that have 0.35 (Figure 10a). The normalized RMSD (nRMSD) is also lower for the XPOL estimates. For the stratiform precipitation event of 12–13 August 2014 (Figure 10b), the XPOL estimates reveal higher correlations (~0.85) and lower RMSD values (close to 1.5). The C-band estimates also exhibit higher correlation (close to 0.7) and relative RMSD close to 1.8. The XPOL precipitation fields had standard deviations close to the reference rainfall fields, while the C-band precipitation estimates greatly underestimate the reference rainfall variability.

3.3. Hydrologic Evaluation of Radar-Rainfall Estimates

The potential benefit in using short-range X-band dual-polarization high-resolution observations for flood modeling applications in mountainous basins is examined by comparing the different radar and gauge-based hydrologic simulations, with the latter used as a reference throughout the hydrologic analysis. Runoff response in the study basin is simulated using a spatially distributed hydrologic model forced with rainfall estimates from the gauges, the operational radar, and the XPOL. The following sections discuss the selected runoff events, the hydrologic modeling setup and the analysis of simulated runoff.

3.3.1. Runoff Events

The main volume of annual runoff in Saldur basin is generated between May and September, which is typical of snow dominated glaciated catchments. According to the tracer-based analysis of [50,66], the main runoff-contributing sources in the basin are as follows. From May to June, the runoff generation is predominantly attributed to melting of snow that has been accumulated during the cold season (late fall and winter). During mid-summer (July to early August), the runoff is largely generated by a mixture of snowmelt and glacier-melt and for the rest of the summer until September, glacier melt contributes to streamflow. At the same time, the precipitation events dominated by liquid phase (i.e., rainfall) translate in high direct runoff manifested as distinct spikes in runoff time series (Figure 11).

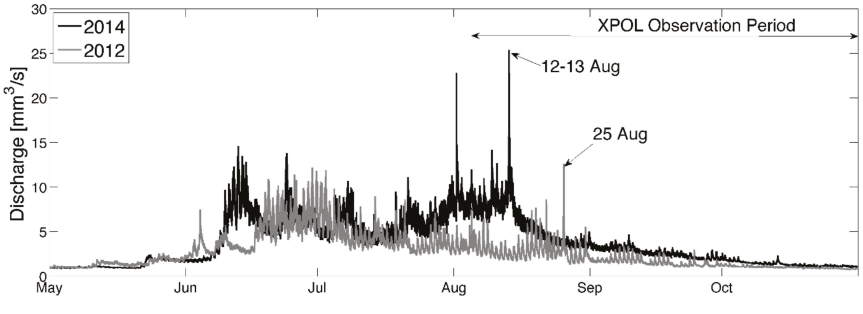


Figure 11. The time series of observed and discharge at the outlet of Saldur basin. The XPOL observation period and the two events analyzed are noted on the graph.

Since the focus of this study is on rainfall-runoff events, only the rainfall events associated with a significant signal in runoff response were selected for hydrologic analysis. From the list of events observed with XPOL (Table 1), only two events resulted in a considerable runoff response for the Saldur basin, which also corresponded to the annual max flow events of 2012 and 2014, respectively. The first event was during 25 August 2012 and was associated with an observed peak flow of ~12 m^3 s^{-1}

(at the basin outlet) and a gauge-based total basin-average rainfall of 21.2 mm. The second event on 13 August 2014 corresponded to an observed peak flow of ~25 m^3 s^{-1} (Figure 11) and was associated with a 32.8 mm total basin-average rainfall. Note that those peak flow values correspond to the minimum and maximum of the annual max flows recorded within a 7-year record (2009–2014) of available discharge observations, respectively (stream gauges).

Based on the air temperatures of the Matsch LTER network, both events happened during a warm summer period. The snow limit was quite high (between 4500 and 3100 for the first one, between 3400 and 3100 for the second one), and therefore could be assumed that most of the basin was contributing by having liquid precipitation. On the basis of previous studies on the catchment, which analyzed the water sourced distribution along the season [70] and for specific events [51] we can estimate that in August roughly 50% of the base flow is due to glacier-melt, the snowmelt is negligible and the peak runoff is due to liquid precipitation.

3.3.2. Hydrologic Model Setup

The hydrologic simulations were carried out with a simple spatially distributed hydrologic model (KLEM—Kinematic Local Excess Model). The model is suitable for event-based rainfall-runoff analysis and has been used in several past studies in similar mountainous regions of northern Italy [71–75]. Model setup requires spatially distributed information of land surface properties (topography, soil type, and land use/cover) that were made available by the Hydrologic Office of the Autonomous Province of Bolzano. The model's spatial resolution is dictated by the resolution of the digital elevation model, which in our case is 30 m. Runoff generation modelling is based on the Soil Conservation Service—Curve Number (SCS–CN) procedure [76]. Representation of runoff routing follows a simple scheme that distinguishes the drainage paths between the channel and hill slopes and assumes two spatially invariant velocities for hillslope and channel grids respectively [77]. The channel and hillslope grids are differentiated based on a channelization support area, which is considered constant at the sub-basin scale. For more information on KLEM model, the interested reader is referred to Reference [78] and references therein.

The model was calibrated to determine the required parameters that include the channelization support area, the channel and hillslope velocity, the parameter of infiltration storativity used in the SCS—CN procedure, and the parameter required for the specification of the initial abstraction [75]. Additionally, the identification of the appropriate antecedent moisture condition group required for CN calculation was also considered in the calibration exercise. Calibration of parameters was carried out manually for the rain gauge (reference) based simulations with the objective to achieve a realistic representation of the observed hydrograph shape and peak. Once calibrated, the model parameters were used consistently for the simulations based on XPOL and C-band rainfall estimates.

It is important to acknowledge that in a glaciated catchment, as our study basin, certain hydrological processes taking place such as glacier melting are not resolved by a relatively simple model like KLEM. However, implicitly accounting for their contribution (e.g., match antecedent flow conditions by proper initialization) permits the use of such a model to realistically simulate rainfall-to-runoff transformation at event scale. Given that the objective here is to examine the impact of differences in event rainfall (magnitude and spatial pattern) to runoff response, we consider that the application of KLEM is adequate for this purpose.

3.3.3. Simulation Results

Simulated runoff and basin-average rainfall, for the two events, are shown in Figure 12 for all the rainfall input scenarios examined (rain gauges, XPOL and C-band radars). The relative errors in the peak and total runoffs are summarized in Table 4.

Table 4. The basin-average rainfall accumulation and the relative difference in runoff peak and volume between the rain gauge and radar-based simulations for the two events of 25 August 2012 and 12–13 August 2014.

	25 August 2012			12–13 August 2014		
	Total Rainfall (mm)	**Rel. Diff. in Peak Runoff (%)**	**Rel. Diff. in Runoff Volume (%)**	**Total Rainfall (mm)**	**Rel. Diff. in Peak Flow (%)**	**Rel. Diff. in Flow Volume (%)**
Rain gauge	21.2	-	-	32.8	-	-
XPOL-200	21.8	33	4.4	36.4	−37	−12.7
XPOL-1000	21.7	27.8	3.8	35.8	−37	−13
C-band	6	−62.4	−35.7	8.8	−70	−21.7

The observed discharge time series are superimposed (Figure 12) to demonstrate the effectiveness of the hydrologic model's calibration procedure, highlighting a good agreement in terms of peak runoff estimation. Comparison against the reference (control) runoff simulation is used in this study to evaluate the impact of radar-rainfall estimation uncertainty in rainfall-to-runoff transformation. One of the most distinct features revealed from the hydrologic simulations is that the C-band based simulated hydrographs showed almost no peak signature, revealing a dramatic underestimation of event flow peak in both cases examined. On the contrary, the XPOL-based simulations resulted in a clear flood hydrograph shape that quite well followed the dynamics of the reference-based simulations. However, it should be noted that XPOL-based simulations were associated with a considerable difference in the estimated hydrograph peak (27–37% with respect to gauge-based simulations) and inconsistent direction of the bias (i.e., underestimation or overestimation). Interestingly, comparison of the basin-average precipitation (Figure 12a,b) shows that the magnitude and dynamics are well captured by the XPOL estimates, which is why the relative difference in total runoff volume remains low (4–13%). This suggests that higher differences in the peak flow originate from differences in rainfall spatial pattern, which are further magnified due to the highly nonlinear propagation of errors from rainfall to peak flow.

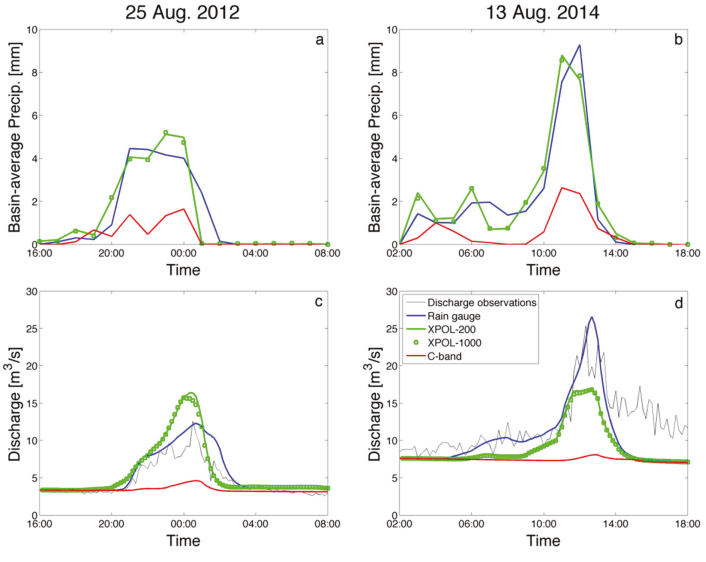


Figure 12. The time series of the basin: the average precipitation (**top**) and simulated discharge (**bottom**) at Saldur outlet (WG1) during 25 August (**left**) and 13 August (**right**), according to rain gauges (blue), XPOL-200 m (green), XPOL-1 km (green squares) and C-band (red) estimates. Available discharge measurements (thin black) are also shown.

The comparison of XPOL results at the two different grid resolutions (200 and 1000 m) shows that difference in pixel resolution did not have an important impact on the simulations of basin's hydrologic response. This is somewhat expected, given previous research on the topic, which has shown that for the ratio of pixel resolution and basin scale considered here, no significant difference is expected in runoff response [11,79]. Overall, the findings from the hydrologic simulations demonstrate the potential benefit of using observations from locally deployed X-band dual-polarization radar for hydrological applications in mountainous areas. The systematic and severe underestimation in radar rainfall estimates from the regional C-band radar does not permit meaningful hydrologic analysis. This is also a clear indication of the potential limitations for using regional radars for developing flood early warning systems in remote complex terrain regions.

4. Conclusions

This study used data from a field experiment in the northeastern Italian Alps to demonstrate quantitative advantages from locally deployed dual-polarization X-band radar (XPOL) observations, relative to operational gauge-adjusted C-band radar-rainfall estimates, in terms of measuring rainfall over mountainous areas and the associated impacts in simulating runoff in small basins. The XPOL radar data were corrected for effects due to rain-path attenuation and the complex terrain (i.e., beam blockage, and VPR correction in areas with bright band and mixed phase contamination). For accurate calibration of radar parameters (Z_H and Z_{DR}), we used measurements from two disdrometers located at a site close to the radar with no intermediate obstacles. In addition, a three-dimensional model of the radar beam was applied to exclude highly occluded areas or to correct reflectivity from areas with minor occlusion (e.g., 3 dB reduction of Z_H). Then polarimetric rainfall estimates were reported in two different spatial (200 and 1000 m) resolutions.

Ten rain events, measured by a locally deployed XPOL radar, dense in situ stations (rain gauges, disdrometers) and two regional operational radars, were selected to represent the case studies for this work. The intensity of the events was low to moderate, which did not induce significant flash flooding in the study area, but, according to the discharge observations, two of these events resulted in considerable direct runoff events. For the evaluation of radar-rainfall estimates, we selected fourteen different validation sites representing a large elevation range between 700 and 3500 m, which is not found in similar studies in the literature.

The comparison of the two XPOL spatial aggregation estimates with the in-situ gauges in both high and low elevations gave high correlations (0.60–0.99). The bias ratio for conditional (>0.1 mm h^{-1}) and unconditional gauge rainfall varied between 20% and 53% overestimation. In two cases the total XPOL rainfall estimates exhibited two to three times larger precipitation accumulations than reference. No significant effect in the estimation accuracy was noted due to spatial resolution. The overall statistics showed better comparisons with the rain gauge measurements at high elevations (i.e., within the Saldur basin) and the XPOL rainfall estimation algorithm gave almost unbiased estimates for rainfall rates greater than 2 mm h^{-1}. On the other hand, both the Austrian and Swiss operational radar network rainfall products gave lower correlations (0.45–0.85) against the in-situ gauges and high underestimation biases (from 0.21 up to 0.91 for the low elevation gauge groups, and from 0.24 up to 0.97 for the high elevation gauge groups).

Inspection of rainfall accumulation maps from the two high flow rainfall events showed that the spatial representation and magnitude of rainfall was significantly underestimated by the operational radar-rainfall products. Runoff simulations based on XPOL estimates were close to the reference (rain gauge) simulations in terms of runoff volume (relative difference <13%), but the peak runoff deviated by a larger degree (~30% relative difference). Simulations obtained from the operational radar rainfall products did not result in a notable peak flow. These results highlight the limitations of long-range radar observations for the flood modeling of small mountainous basins, and the potential from using locally deployed X-band radars as gap filling observing systems, especially in mountain regions, where beam blockage limits the effectiveness of traditional long-range C-band radars.

Even though the sample size associated with ten events was good enough to generate bulk error statistics in terms of precipitation estimation (mainly due to the number of in situ gauges), this study lacks a comprehensive representation of storm types, including rainfall intensities and high impact flood events. Therefore, it remains to be demonstrated how significant the described improvements are for different flood event severities and watershed characteristics, based on field experiments representing varying hydro-climatic regimes.

Author Contributions: M.N.A., J.K., M.B., E.N.A., E.M. and G.B. conceived and designed the experiments; M.N.A. and J.K. performed the XPOL and disdrometer experiments; M.B., E.M., U.T. and G.B. performed the ground validation rain gauge and discharge data and contribute on providing the Valluga and MeteoSwiss C-band radar data; M.N.A., J.K. and F.M. analyzed the precipitation data; E.I.N. carried out the hydrologic simulations and corresponding analysis and contributed the relevant text of the manuscript. E.N.A. contributed on the final revision of the paper; M.N.A. wrote the paper.

Funding: The research leading to these results has received part of the funding from the M. N. Anagnostou Postdoctoral Fellowship under the Action «Supporting Postdoctoral Researchers» of the Operational Program "Education and Lifelong Learning" (Action's Beneficiary: General Secretariat for Research and Technology) and is co-financed by the European Social Fund (ESF) and the Greek State, under the Grand Agreement PE10 (975) HYDRO-X, coordinated by the National Observatory of Athens, Greece. In addition, EURAC activities were partially supported by the research projects "EMERGE: Retreating glaciers and emerging ecosystems in the Southern Alps" (Erich-Ritter-und Herzog-Sellenberg-Stiftung im Stifterverband für die Deutsche Wissenschaft) and "HydroAlp", financed by Autonomous Province of Bolzano, Ripartizione Diritto allo Studio, Università e Ricerca Scientifica.

Acknowledgments: The authors would like to thank Urs Germann and Ioannis Sideris from MeteoSwiss for providing the Swiss Radar Network rainfall observations, and Andreas Schaffhauser (Austro Control Österreichische Gesellschaft, Austria) for providing the Valluga radar rainfall observations. Matteo Vischi and Pierpaolo Macconi (Ripartizione Opere Idrauliche, Provincia Autonoma di Bolzano—Italy) are thanked for the support in the installation of the XPOL radar. The study site is part of the LTSER Platform Matsch | Mazia, a member of the national and international Long-Term Ecological Research Networks (LTER-Italy, LTER-Europe and ILTER). Michael Engel and Francesco Comiti of the University of Bolzano are thanked for providing discharge and rainfall observations. G. Niedrist of EURAC is thanked for his work in maintaining the meteorological stations. Nikolopoulos acknowledges support from EU FP7 Marie Curie Actions IEF (project PIEF-GA-2011-302720).

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

References

1. Petley, D. Global patterns of loss of life from landslides. *Geology* **2012**, *40*, 927–930. [CrossRef]
2. Dowling, C.; Santi, P. Debris flows and their toll on human life: A global analysis of debris flow fatalities from 1950 to 2011. *Nat. Hazards* **2014**, *71*, 203–227. [CrossRef]
3. Borga, M.; Stoffel, M.; Marchi, L.; Marra, F.; Jacob, M. Hydrogeomorphic response to extreme rainfall in headwater systems: Flash floods and debris flows. *J. Hydrol.* **2014**, *518*, 194–205. [CrossRef]
4. Westra, S.; Fowler, H.J.; Evans, J.P.; Alexander, L.V.; Berg, P.; Johnson, F.; Kendon, E.J.; Lenderink, G.; Roberts, N.M. Future changes to the intensity and frequency of short-duration extreme rainfall. *Rev. Geophys.* **2014**, *52*, 522–555. [CrossRef]
5. Molnar, P.; Fatichi, S.; Gaal, L.; Szolgay, J.; Burlando, P. Storm type effects on super Clausius-Clapeyron scaling of intense rainstorm properties with air temperature. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 1753–1766. [CrossRef]
6. UNISDR 2009. Terminology for Disaster Risk Reduction. Available online: http://www.unisdr.org/we/inform/terminology (accessed on 11 January 2013).
7. European Commission. *Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the Assessment and Management of Flood Risks*; European Commission: Brussels, Belgium, 2007.
8. Alfieri, L.; Salamon, P.; Pappenberger, F.; Wetterhall, F.; Thielen, J. Operational early warning systems for water-related hazards in Europe. *Environ. Sci. Policy* **2012**, *21*, 35–49. [CrossRef]
9. Liechti, K.; Zappa, M.; Fundel, F.; Germann, U. Probabilistic evaluation of ensemble discharge nowcasts in two nested Alpine basins prone to flash floods. *Hydrol. Process.* **2013**, *27*, 5–17. [CrossRef]
10. Zoccatelli, D.; Borga, M.; Zanon, F.; Antonescu, B.; Stancalie, G. Which Rainfall Spatial Information for Flash Flood Response Modelling? A Numerical Investigation Based on Data from the Carpathian Range, Romania. *J. Hydrol.* **2010**, *394*, 148–161. [CrossRef]

11. Nikolopoulos, E.I.; Anagnostou, E.N.; Borga, M.; Vivoni, E.R.; Papadopoulos, A. Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability. *J. Hydrol.* **2011**, *402*, 165–178. [CrossRef]
12. Lobligeois, F.; Andréassian, V.; Perrin, C.; Tabary, P.; Loumagne, C. When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 575–594. [CrossRef]
13. Nikolopoulos, E.I.; Crema, S.; Marchi, L.; Marra, F.; Guzzetti, F.; Borga, M. Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence. *Geomorphology* **2014**, *221*, 286–297. [CrossRef]
14. Marra, F.; Nikolopoulos, E.I.; Creutin, J.D.; Borga, M.; Creutin, J.D. Radar rainfall estimation for the identification of debris-flow occurrence thresholds. *J. Hydrol.* **2014**, *519*, 1607–1619. [CrossRef]
15. Barros, A.P.; Joshi, M.; Putkonen, J.; Burbank, D.W. A study of the 1999 monsoon rainfall in a mountainous region in central Nepal using TRMM products and rain gauge observations. *Geophys. Res. Lett.* **2000**, *27*, 3683–3686. [CrossRef]
16. 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]
17. Lewis, H.W.; Harrison, D.L. Assessment of radar data quality in upland catchments. *Meteorol. Appl.* **2007**, *14*, 441–454. [CrossRef]
18. Zhang, J.; Qi, Y.; Kingsmill, D.; Howard, K. Radar-Based Quantitative Precipitation Estimation for the Cool Season in Complex Terrain: Case Studies from the NOAA Hydrometeorology Testbed. *J. Hydrometeorol.* **2012**, *13*, 1836–1854. [CrossRef]
19. White, A.B.; Neiman, P.J.; Ralph, F.M.; Kingsmill, D.E.; Persson, P.O.G. Coastal orographic rainfall processes observed by radar during the California landfalling jets experiment. *J. Hydrometeorol.* **2003**, *4*, 264–282. [CrossRef]
20. Chen, S.; Gourley, J.J.; Hong, Y.; Kirstetter, P.E.; Zhang, J.; Howard, K.; Flamig, Z.L.; Hu, J.; Qi, Y. Evaluation and Uncertainty Estimation of NOAA/NSSL Next-Generation National Mosaic Quantitative Precipitation Estimation Product (Q2) over the Continental United States. *J. Hydrometeorol.* **2013**, *14*, 1308–1322. [CrossRef]
21. Smith, J.A.; Miller, A.J.; Baeck, M.L.; Nelson, P.A.; Fisher, G.T.; Meierdiercks, K.L. Extraordinary flood response of a small urban watershed to short-duration convective rainfall. *J. Hydrometeorol.* **2005**, *6*, 599–617. [CrossRef]
22. Anagnostou, M.N.; Kalogiros, J.; Anagnostou, E.N.; Papadopoulos, A. Experimental results on rainfall estimation in complex terrain with a mobile X-band polarimetric radar. *Atmos. Res.* **2009**, *94*, 579–595. [CrossRef]
23. Anagnostou, M.N.; Kalogiros, J.; Anagnostou, E.N.; Tarolli, M.; Papadopoulos, A.; Borga, M. Performance evaluation of high-resolution rainfall estimation by X-band dual-polarization radar for flash flood applications in mountainous basin. *J. Hydrol.* **2010**, *394*, 4–16. [CrossRef]
24. Wang, Y.; Chandrasekar, V. Quantitative precipitation estimation in the CASA X-band dual-polarization radar network. *J. Atmos. Ocean. Technol.* **2010**, *27*, 1665–1676. [CrossRef]
25. Shakti, P.C.; Maki, M.; Shimizu, S.; Maesaka, T.; Kim, D.; Lee, D.; Iida, H. Correction of Reflectivity in the Presence of Partial Beam Blockage over a Mountainous Region Using X-Band Dual Polarization Radar. *J. Hydrometeorol.* **2013**, *14*, 744–764.
26. Matrosov, S.Y.; Cifelli, R.; Gochis, D. Measurements of Heavy Convective Rainfall in the Presence of Hail in Flood-Prone Areas Using an X-Band Polarimetric Radar. *J. Appl. Meteorol. Climatol.* **2013**, *52*, 395–407. [CrossRef]
27. Koffi, A.K.; Gosset, M.; Zahiri, E.-P.; Ochou, A.D.; Kacou, M.; Cazenave, F.; Assamoi, P. Evaluation of X-band polarimetric radar estimation of rainfall and rain drop size distribution parameters in West Africa. *Atmos. Res.* **2014**, *143*, 438–461. [CrossRef]
28. 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]
29. Marra, F.; Morin, E. Autocorrelation structure of convective rainfall in semiarid-arid climate derived from high-resolution X-band radar estimates. *Atmos. Res.* **2018**, *200*, 126–138. [CrossRef]
30. Chandrasekar, V.; Wang, Y.; Maki, M.; Nakane, K. Flood Monitoring using X-band Dual-polarization Radar Network. In Proceedings of the 11th Plinius Conference on Mediterranean Storms, Barcelona, Spain, 7–10 September 2009.

31. Matrosov, S.Y.; Clark, A.; Martner, B.E.; Tokay, A. X-band polarimetric radar measurements of rainfall. *J. Appl. Meteorol.* **2002**, *41*, 941–952. [CrossRef]
32. Anagnostou, E.N.; Anagnostou, M.N.; Krajewski, W.F.; Kruger, A.; Miriovsky, B.J. High-resolution rainfall estimation from X-band polarimetric radar measurements. *J. Hydrometeorol.* **2004**, *5*, 110–128. [CrossRef]
33. Park, S.-G.; Maki, M.; Iwanami, K.; Bringi, V.N.; Chandrasekar, V. Correction of Radar Reflectivity and Differential Reflectivity for Rain Attenuation at X Band. Part I: Theoretical and Empirical Basis. *J. Atmos. Ocean. Technol.* **2005**, *22*, 1621–1632. [CrossRef]
34. Kim, D.-S.; Maki, M.; Lee, D.-I. Retrieval of three dimensional raindrop size distribution using X-band polarimetric radar. *J. Atmos. Ocean. Technol.* **2010**, *27*, 1265–1285. [CrossRef]
35. Kalogiros, J.; Anagnostou, M.N.; Anagnostou, E.N.; Montopoli, M.; Picciotti, E.; Marzano, F.S. Correction of Polarimetric Radar Reflectivity Measurements and Rainfall Estimates for Apparent Vertical Profile in Stratiform Rain. *J. Appl. Meteorol. Climatol.* **2013**, *52*, 1170–1186. [CrossRef]
36. Kalogiros, J.; Anagnostou, M.N.; Anagnostou, E.N.; Montopoli, M.; Picciotti, E.; Marzano, F.S. Optimum estimation of rain microphysical parameters using X-band dual-polarization radar observables. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3063–3076. [CrossRef]
37. Anagnostou, M.N.; Kalogiros, J.; Marzano, F.S.; Anagnostou, E.N.; Montopoli, M.; Picciotti, E. Performance evaluation of a new dual-polarization microphysical algorithm based on long-term X-band radar and disdrometer observations. *J. Hydrometeorol.* **2013**, *14*, 560–576. [CrossRef]
38. Lim, S.; Cifelli, R.; Chandrasekar, V.; Matrosov, S.Y. Precipitation Classification and Quantification Using X-Band Dual-Polarization Weather Radar: Application in the Hydrometeorology Testbed. *J. Atmos. Ocean. Technol.* **2013**, *30*, 2108–2120. [CrossRef]
39. Chang, W.-Y.; Vivekanandan, J.; Wang, T.-C.C. Estimation of X-Band Polarimetric Radar Attenuation and Measurement Uncertainty Using a Variational Method. *J. Appl. Meteorol. Climatol.* **2014**, *53*, 1099–1119. [CrossRef]
40. Thurai, M.; Mishra, K.V.; Bringi, V.N.; Krajewski, W.F. Initial Results of a New Composite-Weighted Algorithm for Dual-Polarized X-band Rainfall Estimation. *J. Hydrometeorol.* **2017**, *18*, 1081–1100. [CrossRef]
41. Matrosov, S.Y.; Kingsmill, D.E.; Martner, B.E.; Ralph, F.M. The utility of X-band polarimetric radar for quantitative estimates of rainfall parameters. *J. Hydrometeorol.* **2005**, *6*, 248–262. [CrossRef]
42. Chen, H.; Chandrasekar, V. The Quantitative Precipitation Estimation System for Dallas-Fort Worth (DFW) Urban Remote Sensing Network. *J. Hydrol.* **2015**, *531 Pt 2*, 259–271. [CrossRef]
43. Yang, W.-Y.; Ni, G.-H.; Qi, Y.-C.; Hong, Y.; Sun, T. Exploring the potential of utilizing high resolution X-band radar for urban rainfall estimation. *Atmos. Meas. Tech. Discuss.* **2016**. [CrossRef]
44. Chandrasekar, V.; Wang, Y.; Chen, H. The CASA quantitative precipitation estimation system: A five year validation study. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 2811–2820. [CrossRef]
45. 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.* **2018**, *96A*. [CrossRef]
46. Mishra, K.V.; Krajewski, W.F.; Goska, R.; Ceynar, D.; Seo, B.-C.; Kruger, A.; Niemeier, J.J.; Galvez, M.B.; Thurai, M.; Bringi, V.N.; et al. Deployment and Performance Analyses of High-Resolution Iowa XPOL Radar System during the NASA IFloodS Campaign. *J. Hydrometeorol.* **2016**, *17*, 455–479. [CrossRef]
47. Lim, S.; Lee, D.-R.; Cifelli, R.; Hwang, S.H. Quantitative precipitation estimation for an X-band dual-polarization radar in the complex mountainous terrain. *KSCE J. Civ. Eng.* **2014**, *18*, 1548–1553. [CrossRef]
48. Paschalis, A.; Fatichi, S.; Molnar, P.; Rimkus, S.; Burlando, P. On the effects of small scale space-time variability of rainfall on basin flood response. *J. Hydrol.* **2014**, *514*, 313–327. [CrossRef]
49. Comiti, F.; Marchi, L.; Macconi, P.; Arattano, M.; Bertoldi, G.; Borga, M.; Brardinoni, F.; Cavalli, M.; D'Agostino, V.; Penna, D.; et al. A new monitoring station for debris flows in the European Alps: First observations in the Gadria basin. *Nat. Hazards* **2014**, *73*, 1175–1198. [CrossRef]
50. Mao, L.; Dell'Agnese, A.; Huincache, C.; Penna, D.; Engel, M.; Niedrist, G.; Comiti, F. Bedload hysteresis in a glacier-fed mountain river. *Earth Surf. Process. Landf.* **2015**, *39*, 964–976. [CrossRef]
51. Engel, M.; Penna, D.; Bertoldi, G.; Dell'Agnese, A.; Soulsby, C.; Comiti, F. Identifying run-off contributions during melt-induced run-off events in a glacierized alpine catchment. *Hydrol. Process.* **2016**, *30*, 343–364. [CrossRef]

52. Galos, S.; Kaser, G. *The Mass Balance of Matscherferner 2012/13, Project Report*; University of Innsbruck: Innsbruck, Austria, 2014.
53. Engel, M.; Notarnicola, C.; Endrizzi, S.; Bertoldi, G. Snow model sensitivity analysis to understand spatial and temporal snow dynamics in a high-elevation catchment. *Hydrol. Process.* **2017**, *31*, 4151–4168. [CrossRef]
54. Penna, D.; Engel, M.; Mao, L.; Dell'Agnese, A.; Bertoldi, G.; Comiti, F. Tracer-based analysis of spatial and temporal variation of water sources in a glacierized catchment. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 5271–5288. [CrossRef]
55. Bertoldi, G.; Della Chiesa, S.; Notarnicola, C.; Pasolli, L.; Niedrist, G.; Tappeiner, U. Estimation of soil moisture patterns in mountain grasslands by means of SAR RADARSAT 2 images and hydrological modelling. *J. Hydrol.* **2014**, *516*, 245–257. [CrossRef]
56. Mair, E.; Leitinger, G.; della Chiesa, S.; Niedrist, G.; Tappeiner, U.; Bertoldi, G. A simple method to combine snow height and meteorological observations to estimate winter precipitation at sub-daily resolution. *Hydrol. Sci. J.* **2015**, *61*, 2050–2060. [CrossRef]
57. Schaffhauser, A.; Auer, M.; Kann, A. Weather Radar Observations at high altitude is it worth the effort? In Proceedings of the 6th European Conference on Radar in Meteorology and Hydrology (ERAD), Sibiu, Romania, 6–10 September 2010.
58. Paulitsch, H.; Teschl, F.; Randeu, W.L. Dual-polarization C-band weather radar algorithms for rain rate estimation and hydrometeor classification in an alpine region. *Adv. Geosci.* **2009**, *20*, 3–8. [CrossRef]
59. Paulitsch, H.; Teschl, F.; Randeu, W.L. Preliminary evaluation of polarimetric parameters from a new dual-polarization C-band weather radar in an alpine region. *Adv. Geosci.* **2010**, *25*, 111–117. [CrossRef]
60. Germann, U.; Galli, G.; Boscacci, M.; Bolliger, M. Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.* **2006**, *132*, 1669–1692. [CrossRef]
61. Sideris, I.; Gabella, M.; Sassi, M.; Germann, U. Real-time spatiotemporal merging of radar and raingauge precipitation measurements in Switzerland. In Proceedings of the 9th International Workshop on Precipitation in Urban Areas, Urban Challenges in Rainfall Analysis, St. Moritz, Switzerland, 6–9 December 2012.
62. Sideris, I.; Gabella, M.; Erdin, R.; Germann, U. Real-time radar-raingauge merging using spatio-temporal co-kriging with external drift in the Alpine terrain of Switzerland. *Q. J. R. Meteorol. Soc.* **2014**, *140*, 1097–1111. [CrossRef]
63. Sideris, I.; Gabella, M.; Sassi, M.; Germann, U. The CombiPrecip experience: Development and operation of a real-time radar-raingauge combination scheme in Switzerland. In Proceedings of the 9th Weather Radar and Hydrology (WRaH) International Symposium, Washington, DC, USA, 7–9 April 2014.
64. Kalogiros, J.; Anagnostou, M.N.; Anagnostou, E.N.; Montopoli, M.; Picciotti, E.; Marzano, F.S. Evaluation of a new Polarimetric Algorithm for Rain-Path Attenuation Correction of X-Band Radar Observations Against Disdrometer Data. *IEEE Geosci. Remote Sens. Lett.* **2014**, *52*, 1369–1380. [CrossRef]
65. 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]
66. Gorgucci, E.; Chandrasekar, V.; Baldini, L. Correction of X-band radar observation for propagation effects based on the self-consistency principle. *J. Atmos. Ocean. Technol.* **2006**, *23*, 1668–1681. [CrossRef]
67. Matrosov, S.Y.; Kurt, C.; Kingsmill, A.; David, E. A polarimetric radar approach to identify rain. Melting-layer and snow regions for applying corrections to vertical profiles of reflectivity. *J. Appl. Meteorol. Climatol.* **2007**, *46*, 154–166. [CrossRef]
68. Borga, M.; Anagnostou, E.N.; Frank, E. On the use of real-time radar rainfall estimates for flood prediction in mountainous basins. *J. Geophys. Res. Atmos.* **2000**, *105*, 2269–2280. [CrossRef]
69. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* **2001**, *106*, 7183–7192. [CrossRef]
70. Penna, D.; Engel, M.; Bertoldi, G.; Comiti, F. Towards a tracer-based conceptualization of meltwater dynamics and streamflow response in a glacierized catchment. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 23–41. [CrossRef]
71. Sangati, M.; Borga, M.; Rabuffetti, D.; Bechini, R. Influence of rainfall and soil properties spatial aggregation on extreme flash flood response modelling: An evaluation based on the Sesia river basin, North Western Italy. *Adv. Water Res.* **2009**, *32*, 1090–1106. [CrossRef]

72. Zoccatelli, D.; Borga, M.; Viglione, A.; Chirico, G.B.; Blöschl, G. Spatial moments of catchment rainfall: Rainfall spatial organisation, basin morphology, and flood response. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3767–3783. [CrossRef]
73. Nikolopoulos, E.I.; Borga, M.; Zoccatelli, D.; Anagnostou, E.N. Catchment-Scale Storm Velocity: Quantification, Scale Dependence and Effect on Flood Response. *Hydrol. Sci. J.* **2014**, *59*, 1363–1376. [CrossRef]
74. Bartsotas, N.S.; Nikolopoulos, E.I.; Anagnostou, E.N.; Solomos, S.; Kallos, G. Moving Toward Subkilometer Modeling Grid Spacings: Impacts on Atmospheric and Hydrological Simulations of Extreme Flash Flood-Inducing Storms. *J. Hydrol.* **2017**, *18*, 209–226. [CrossRef]
75. Destro, E.; Amponsah, W.; Nikolopoulos, E.I.; Marchi, L.; Marra, F.; Zoccatelli, D.; Borga, M. Coupled prediction of flash flood response and debris flow occurrence: Application on an alpine extreme flood event. *J. Hydrol.* **2018**, *558*, 225–237. [CrossRef]
76. U.S. Department of Agriculture. *Urban Hydrology for Small Watersheds*; Technical Release; U.S. Department of Agriculture: Washington, DC, USA; Volume 55, p. 164.
77. Da Ros, D.; Borga, M. Use of digital elevation model data for the derivation of the geomorphological instantaneous unit hydrograph. *Hydrol. Process.* **1997**, *11*, 13–33. [CrossRef]
78. Borga, M.; Boscolo, P.; Zanon, F.; Sangati, M. Hydrometeorological Analysis of the 29 August 2003 Flash Flood in the Eastern Italian Alps. *J. Hydrometeorol.* **2007**, *8*, 1049–1067. [CrossRef]
79. Sangati, M.; Borga, M. Influence of rainfall spatial resolution on flash flood modeling. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 575–584. [CrossRef]

remote sensing

MDPI

Article

Observed High-Latitude Precipitation Amount and Pattern and CMIP5 Model Projections

Ali Behrangi [1,*] and Mark Richardson [2,3]

1 Department of Hydrology and Atmospheric Sciences, University of Arizona, 1133 E. James E Rogers Way, Harshbarger, Tucson, AZ 85721, USA
2 Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, MS 233-300, Pasadena, CA 91109, USA; markr@jpl.caltech.edu
3 Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA 90095, USA
* Correspondence: Behrangi@email.arizona.edu; Tel.: +1-520-626-6639

Received: 3 August 2018; Accepted: 24 September 2018; Published: 1 October 2018

Abstract: Utilizing reanalysis and high sensitivity W-band radar observations from CloudSat, this study assesses simulated high-latitude (55–82.5°) precipitation and its future changes under the RCP8.5 global warming scenario. A subset of models was selected based on the smallest discrepancy relative to CloudSat and ERA-I reanalysis using a combined ranking for bias and spatial root mean square error (RMSE). After accounting for uncertainties introduced by internal variability due to CloudSat's limited four year day-night observation period, RMSE provides greater discrimination between the models than a typical mean state bias criterion. Over 1976–2005 to 2071–2100, colder months experience larger fractional modelled precipitation increases than warmer months, and the observation-constrained models generally report a larger response than the full ensemble. For everywhere except the Southern Hemisphere (SH55, for 55–82.5°S) ocean, the selected models show greater warming than the model ensemble while their hydrological sensitivity (fractional precipitation change with temperature) is indistinguishable from the full ensemble relationship. This indicates that local thermodynamic effects explain much of the net high-latitude precipitation change. For the SH ocean, the models that perform best in the present climate show near-median warming but greater precipitation increase, implying a detectable contribution from processes other than local thermodynamic changes. A Taylor diagram analysis of the full CMIP5 ensemble finds that the Northern Hemisphere (NH55) and SH55 land areas follow a "wet get wetter" paradigm. The SH55 land areas show stable spatial correlations between the simulated present and future climate, indicative of small changes in the spatial pattern, but this is not true of NH55 land. This shows changes in the spatial pattern of precipitation changes through time as well as the differences in precipitation between wet and dry regions.

Keywords: precipitation; remote sensing; CloudSat; CMIP; high latitude

1. Introduction

Warming in high latitudes is faster than in lower latitudes partly due to meridional heat transport and positive snow/ice-albedo feedback [1,2]. In the Northern Hemisphere, rapid reductions in snow cover, sea-ice extent [3], and thawing permafrost [4] have been observed and an intensification of the hydrological cycle is expected [5]. Warming has also increased sea level through terrestrial ice and glacier loss, but this can be locally mitigated by an increase in snowfall in regions that remain below freezing even under warming [6,7], meaning that understanding high-latitude precipitation is important for calculations of future global-scale risks.

Climate models are key to understanding future precipitation changes but produce a wide range of values, especially on regional scales [8]. Meanwhile, precipitation observations show large spread, especially in high latitudes (e.g., [9]), which is related to limitations in sensors, retrieval techniques, poor understanding of precipitation microphysics, unknown surface emissivity (i.e., over snow and ice), and difficulties in distinguishing between light rain and cloud [10,11]. In recent years, advanced sensors such as CloudSat (Stephens et al., 2008) have improved our quantification of the amount and distribution of high-latitude precipitation. CloudSat has high sensitivity to detect light rainfall and snowfall, two major types of precipitation in high latitudes. Behrangi et al. [12] performed a comparative analysis between CloudSat total precipitation estimates and other products for the regions 55–82.5°S/N (henceforth "SH55" and "NH55", the polar limit is based on the CloudSat orbit on a 2.5° × 2.5° latitude-longitude grid). The products used were the monthly Global Precipitation Climatology Project (GPCP) V2.3 [13,14], Global Precipitation Climatology Centre (GPCC) full data reanalysis V7 [15], National Centers for Environmental Prediction–Department of Energy Reanalysis 2 (NCEP-DOE R2) [16], ERA-interim [17], and Modern Era Retrospective-Analysis for Research and Applications (MERRA) [18,19]. Mean precipitation and spatial statistics showed that both ERA-I and GPCP generally agreed well with CloudSat, although CloudSat does not have a formal rainfall product over land and so it was not considered robust over NH55 land. One area of disagreement was winter precipitation over northern Eurasia, where GPCP was a high outlier relative to the other non-CloudSat products, and where independent GRACE data did not show a mass increase consistent with GPCP accumulations.

Here our objective is to use the updated understanding of precipitation over SH55 and NH55 land and ocean from Behrangi et al. [12] to assess model performance and determine how current understanding of the amount and spatial distribution of high-latitude precipitation can be used to add insights onto projected future changes in precipitation under global warming. Other studies have assessed how Coupled Model Intercomparison Project, phase 5 (CMIP5) models simulate precipitation (e.g., [20–22]), but here we specifically consider higher latitudes.

One approach to potentially narrow ranges in climate projections is to sub-select models based on how well they simulate properties or processes in the current climate. One example of this approach is that of "emergent constraints" in research related to climate or hydrological sensitivity [23]. With regards to precipitation, Palerme et al. [24] compared CloudSat snowfall with that of CMIP5 models over Antarctica and identified that sea-ice extent is a key predictor of simulated Antarctic snowfall. They also considered a subset of CMIP5 models whose mean Antarctic precipitation was within ±20% of that from the CloudSat 2C-SNOW product. Here we consider land and ocean in both hemispheres, rather than just Antarctica, and assess whether the observation-based spatial pattern of precipitation can be used as an additional constraint to reduce the modelled spread of future high-latitude changes.

The use of spatial distribution of precipitation in principle means that greater information can be extracted from a relatively short period, such as the 4 full years for which CloudSat operated both day and night. Some previous studies (i.e., [25]) have found a dominant role of bias in determining model skill scores. We find that models disagree more strongly in terms of their simulated spatial pattern of precipitation than in terms of their means, relative to the uncertainties introduced over a 4-year period by internal variability. This indicates that additional use of spatial information provides a more robust method for ranking models by performance and determining whether better performing models indicate different future changes.

2. Dataset

Building on Behrangi et al. [12], we select reference observation datasets as follows: (1) CloudSat for rain and snowfall over ocean and precipitation over Antarctica where snowfall accounts for 99% of total precipitation, and (2) ERA-interim (here after ERA-I) over Northern Hemisphere land where rainfall is a major contributor and a reliable CloudSat retrieval is not available. Note that while GPCP

showed similar agreement with CloudSat for Antarctic precipitation, several studies suggest that GPCP over estimates precipitation in parts of Eurasia, which may be related to biases in the correction for gauge undercatch. We selected ERA-I as it was more consistent compared to CloudSat over both NH55 ocean and SH55 land and ocean. Data were mapped onto a 2.5° × 2.5° latitude-longitude grid, which allows sufficient CloudSat samples to stabilize grid-box-level statistics.

2.1. CloudSat

Due to its high sensitivity to light rain and snowfall, which occur frequently in high latitudes, CloudSat has been widely used in analysis of high latitude precipitation [26–29]. This is particularly important due to the sparseness of stations and inherent limitations of other satellite-based precipitation estimates in high latitudes [26,27]. Three CloudSat products are used here: 2C-PRECIP-COLUMN R04 (henceforth "2c-column"; [30]), 2C-RAIN-PROFILE R04 (henceforth "2c-rain") [11], and 2C-SNOW-PROFILE (henceforth "2c-snow") [31] along with the CloudSat auxiliary AMSR-E product CS_AMSRE-AUX (henceforth "AMSR-E"). All data are available at: http://www.cloudsat.cira.colostate.edu. 2c-column provides precipitation occurrence, phase (rain, snow or mixed), and likelihood (certain, probable, possible). We restrict our analysis to "certain" for all phases. Non-certain events tend to be of low intensity and in the case of 2c-snow, including non-certain events only increases mean snowfall rate in the SH55 region by +2.9% and in the NH55 region by +4.5%.

It should be noted that while CloudSat provides a viable source for precipition estimation in high latitudes, it may also face some limitations. In very intense rain the radar signal is saturated and the CloudSat algorithm may provide a lower limit on the rain intensity. This may limit the ability of CloudSat at the higher end of snowfall intensity distribution [32,33]. Signal saturation is often infrequent in high latitudes (especially poleward of latitude 60° [26]), but when it does, coincident AMSR-E estimates from CS_AMSRE-AUX can be used similar to Behrangi et al. [26]. Addition of AMSR-E is particularly effective over ocean where AMSR-E precipition retrievals are based on both emission and scattering signals. Accordingly, the CloudSat estimates used in the present study employs AMSR-E data when signal saturation is noted. CloudSat may also miss shallow precipitation. In CloudSat product, near surface precipitation is estimated at about 1.0 km above the ocean surface (and ~1.5 km over land) to avoid contamination of the reflectivity profile by surface returns [34] that can result in missing shallow precipitation. This limitation has been shown by comparing CloudSat estimates with ground stations and radars (e.g., [32]), although as stated by the authors identifying a reliable ground truth for evaluation of light-snowfall events is difficult. In another study over Antarctica, it was shown that near surface sublimation of snowfall could be large thus CloudSat precipitation estimates at about 1.5 km over land might produce larger rates than what is observed on ground, mitigating part of the missed shallow precipitation by CloudSat [35].

CloudSat products are available from the summer of 2006 but battery problems forced daylight-only operation since April 2011. As in Behrangi et al. [12], we use the complete years 2007–2010 inclusive.

2.2. ERA-Interim

ERA-interim [16] is a European Center for Medium-Range Weather Forecasts (ECMWF) global atmospheric reanalysis that uses a 4D-VAR scheme to assimilate observations from radiosondes, commercial aircraft, and satellites in a numerical model. Precipitation data are not directly incorporated. This study uses daily precipitation from http://apps.ecmwf.int/datasets/ at 2.5° × 2.5° spatial resolution. Previous evaluation has shown good performance of ERA-I in various high-latitude locations. For example, Medley et al. [36] showed that ERA-I's mean snow accumulation was within approximately 1σ of airborne radar observations over Thwaites Glacier in West Antarctica. Other work suggests that ERA-I likely more realistically depicts precipitation changes in Antarctica [37] and high-latitude precipitation amount [12] compared to several other reanalyses. The choice ERA-I over

NH55 land is mainly based on Behrangi et al. [12] which highlighted potential errors such as gauge undercatch issues in GPCP and GPCC, and the lack of reliable CloudSat rain retrievals over land. Nevertheless, we recognize that the use of ERA-I is also not ideal as ERA-I and CMIP5 models could potentially share similar parameterizations.

2.3. CMIP5 Models

We used output from the 36 CMIP5 climate models listed in Table 1. From the middle of the nineteenth century to 2005, the models are forced by known solar output and atmospheric composition, this is the "historical" experiment [38]. From 2005 until the end of the 21st century, the models are forced under a scenario of high emissions called Representative Concentration Pathway 8.5 (RCP8.5, [39]). This choice is not a statement about the likelihood of RCP 8.5 over other potential pathways, but as it provides a forcing that is both credible and large, giving a larger signal-to-noise ratio for interpreting forced response. After assessing the general high-latitude performance of the models we sub-select 5 for each region (SH55 and NH55 land and ocean, plus Greenland) based on their combined performance in terms of spatial RMSE and mean bias relative to the observation-based product. SH55 land corresponds to Antarctica, and Greenland is separated from the rest of NH55 land for two main reasons. Firstly, it is a unique measurement challenge with its perennial ice and snow cover and poor gauge coverage. Secondly, its potential contribution to sea level change is of great interest and is the focus of an active research community. We therefore separate Greenland to avoid cross-contamination of region-specific errors with other NH55 land, and to provide more easily applied information for ice sheet modelers and sea level rise specialists. The selection methodology is described in Section 3.

Table 1. The list of the models used in this study. For each study region the selected models based on Bias, RMSE, or both are labeled by "B", "R", or "X", respectively. ANT and GL are for Antarctica and Greenland.

Models	ANT	NH Ocean	SH Ocean	NH Land	GL	Models	ANT	NH Ocean	SH Ocean	NH Land	GL
NorESM1-M			X			GFDL-ESM2M					X
CMCC-CMS				X	X	bcc-csm1-1-m					
NorESM1-ME	X			X		CNRM-CM5					
MPI-ESM-MR						MRI-CGCM3				X	X
GISS-E2-R						MIROC-ESM					
MPI-ESM-LR						MIROC-ESM-CHEM					
GISS-E2-R-CC						MRI-ESM1					
GFDL-CM3						HadGEM2-CC	X	X		X	X
CSIRO-Mk3-6-0			X			GISS-E2-H		X			
CMCC-CESM						HadGEM2-ES					
CMCC-CM						ACCESS1-0					X
bcc-csm1-1			X			ACCESS1-3	X				
CCSM4		X	X			FGOALS-g2					
CESM1-CAM5	X	X	X			inmcm4					
CESM1-BGC						HadGEM2-AO					
GFDL-ESM2G						MIROC5		X		X	
CanESM2	X					FIO-ESM					
GISS-E2-H-CC						BNU-ESM					

2.4. Other Datasets

The present work also uses precipitation data from GPCP, GPCC, MERRA, and NCEP-DOE R2. GPCP is a merged product using data from gauges over land and from satellite over both land and ocean. We used the latest version of the monthly 2.5° × 2.5° resolution GPCP product (version 2.3; [12]).

GPCC integrates a large pool of station data from various networks, organizations, and additional resources under support of WMO and produces gridded products at a range of spatial resolutions at daily and monthly time scales [14]. Here we used GPCC Full Data Reanalysis version 7.0 at 2.5° × 2.5° resolution. MERRA [18] uses the Goddard Earth Observing System Data Assimilation System version 5 and assimilates observations for the retrospective analyses. Here we used monthly MERRA V5.2 with 0.66° longitude × 0.50° latitude resolution. The NCEP-DOE R2 product [15] is an improved version of the NCEP product and includes fixed errors and updated parameterizations of physical processes. We used monthly data at T62 spatial resolution (~1.875° × 1.875°).

3. Method

Model ensembles such as CMIP5 are often used to characterise potential future climate change, although as an ensemble of opportunity, CMIP5 cannot be used to calculate formal errors. However, investigating its simulation of the present day climate may reveal model-observation discrepancies that indicate poor model performance and it may also be possible to reduce the range of projections through observation-based constraints that exclude poorly performing models.

Here we investigate the CMIP5 ensemble simulation of high-latitude precipitation based on the observation-based product selection justified in Behrangi et al. [12], and then consider whether the present day simulation of high-latitude precipitation is informative regarding future changes within the ensemble. We do so by selecting models that show the smallest discrepancy relative to CloudSat and ERA-I in both the total amount and spatial distribution of high-latitude precipitation. However, the observational data choice, model selection criteria, and the number of models to include in this subset are subjective. Palerme et al. [24] selected models with a mean discrepancy in mean Antarctic precipitation of ±20% versus CloudSat 2C-snow. Here we include spatial information by using centered and area-weighted RMSE as well as bias. All models are ranked separately for each of SH55 and NH55 land and ocean, and for Greenland, and those within the five lowest sum of rankings for bias + RMSE are used in our primary analysis. Table 1 lists all 36 models used along with labels for those which were within the top 5 rankings for bias, RMSE or bias + RMSE. In many cases a model that is selected based on bias is also picked in the RMSE ranking, but this is not always the case. We present results for both the full ensemble and our 5-member subset, and to address potential sensitivities of our method we also show some results when selecting on bias alone.

While CloudSat provides advanced precipitation data to evaluate the models, only 2007–2010 provide full day-night coverage. For all other datasets we use 1986–2005 which raises questions about the stability and reliability of our rankings. We address this by using GPCP as a test of variability introduced by sub-selecting smaller time periods by comparing 4-year statistics with the values inferred over 1990–2010. This comparison includes the CloudSat period and is of the same length as the other periods used. A further advantage of the later selection is to avoid discontinuities early in the GPCP record, where our analysis shows a larger disagreement in spatial pattern between pre-1990 and later years. These larger changes are more likely to be due to discontinuities introduced by changes in data sources rather than real changes in precipitation distribution.

We calculate running 4-year means and RMSEs relative to the 1990–2010 mean pattern for each region and the time series of these statistics are shown in Figure 1. These represent an estimate of the variance introduced due to selecting a four year period compared with the 21-year periods used in our other comparisons.

The standard deviation of the 4-year mean precipitation values provides an estimate of uncertainty in the bias due to the limit of a 4-year CloudSat period whereas the mean value of the RMSE is an estimate of the uncertainty in the RMSE statistic. The variation in the mean precipitation over the oceans and Greenland is similar to the threshold bias used for selecting the top 5 models, indicating that this short time period does not provide a particularly useful constraint on the CMIP5 models. However, the mean RMSE statistic is consistently smaller than that which occurs when comparing CMIP5 simulations with CloudSat. For example, the mean RMSE over SH55 oceans is 0.08 mm day^{-1}, compared with

CMIP5 simulations which range from 0.41 to 0.61 mm day^{-1}. Meanwhile, over NH55 oceans the 4-year CloudSat RMSE is 0.18 mm day^{-1} while the CMIP5 values range from 0.66–0.93 mm day^{-1}. Our use of spatial information through the RMSE therefore allows more consistent discrimination between CMIP5 models despite the short time period of the available CloudSat record. Our use of the additional bias criterion simply ensures that the total precipitation is within a realistic range.

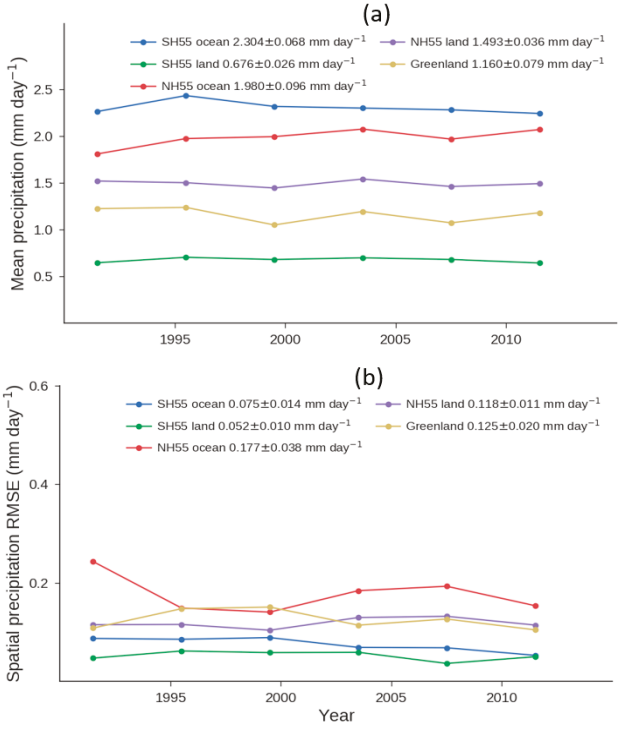


Figure 1. (**a**) mean 4-year average precipitation for each labelled region in GPCP and (**b**) mean 4-year RMSE compared with the 1990–2015 average in GPCP. The points are plotted at the center of each 4-year period.

4. Results

4.1. Mean Precipitation Rate in Observations and Models

Figure 2 shows the time series of total regional precipitation with the CMIP5 5–95% ensemble range in blue and median as bold blue line for years 1850–2100. Observations and reanalyses (GPCP, CloudSat, MERRA, NCEP, ERA-I) values for 2007–2010 (and based on [12]) are shown as the vertical line with the reference observational product (CloudSat over oceans and Antarctica, and ERA-I over NH land) as a black circle. While in NH the models' median fall in the range determined by observations and reanalyses, in SH the models' median exceed this range. Specifically, CloudSat's Antarctic precipitation estimate is below the 5th percentile of CMIP5, as reported in Palerme et al. [24]. While Behrangi et al. [12] showed agreement between CloudSat and GPCP over Antarctica, CloudSat likely underestimates total snowfall, particularly in the Antarctic interior due to ground clutter and missing shallow precipitation [31].

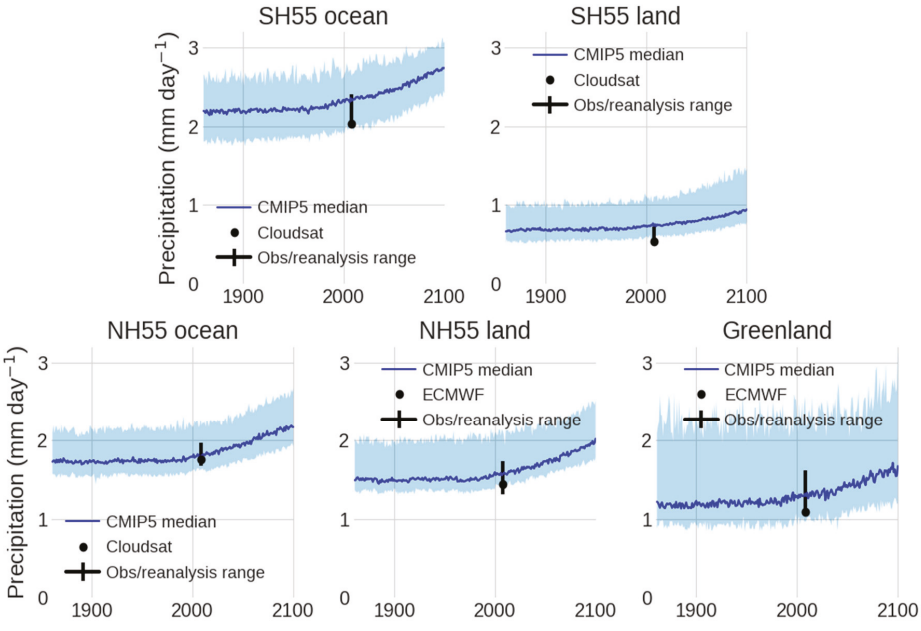


Figure 2. Time series of total regional precipitation with CMIP5 5–95% ensemble range in blue and median as bold blue line for years 1850–2100. Observations and reanalyses (GPCP, CloudSat, MERRA, NCEP, ERA-I) values for 2007–2010 (and based on Behrangi et al., 2016) are shown as the vertical line with the reference observational product (CloudSat over oceans and Antarctica, and ERA-I over NH land) as a black circle.

Figure 3 compares normalized density plots of linear trends in precipitation for models, GPCP, ERA-I, and GPCC. The model ensemble 5–95% range is shown in pink and the median as a red line for 1976–2005. GPCP and GPCC plots are constructed from 1980 to 2009 and are shown in blue and green lines, respectively. GPCC is only available for NH land. The dashed magenta lines show the same properties for the five selected models listed in Table 1. The highest model-observation agreement occurs over NH land, where observation products benefit from the availability of ground stations. The trends are least consistent over SH ocean, where GPCP often shows larger positive slopes in winter and spring, and larger negative slopes in summer and fall. ERA-I trends are generally more consistent with models than GPCP, especially over ocean.

4.2. Future Precipitation Changes

Figure 4 shows maps of annual mean precipitation for 2007–2010 from CloudSat (Figure 4a,b) and ERA-I (Figure 4c,d). It also shows maps of all-model annual mean precipitation for 1976–2005 (Figure 4e,f) and percent precipitation change (by comparing 1971–2100 with 1976–2005) for all (Figure 4g,h) and the selected subset (Figure 4i,j) of models in NH and SH. Figure 4 suggests that while the overall mean precipitation pattern is comparable among the studies products, ERA-I clearly shows lower mean precipitation rate than other products over the Southern Oceans. It can be clearly seen that the selected models suggest a generally larger percentage increase in precipitation than full ensemble average, especially over the Arctic Ocean and Eastern Antarctic ice sheet. The readers are referred to our earlier publication [12] for more detailed comparison between CloudSat precipitation maps and other satellite and reanalysis products. In Figure 5, the fractional regional precipitation changes are plotted separately for each calendar month. The median and 20–80% ensemble range are shown

in blue (for all models) and red (for the subset). Also displayed is the median of the same value for the 5 models selected based on bias alone. Figure 5 suggests that modelled colder months experience larger fractional precipitation increases than warmer months, and that the NH regions see greater increases than those in the SH. The selected models generally report larger fractional increase than the full ensemble, with a larger response than that seen when selecting only based on bias. This is more distinct over the Antarctic and SH ocean where the range is above the full ensemble median across almost all months. Meanwhile, this difference is minimal over Greenland compared to the other regions.

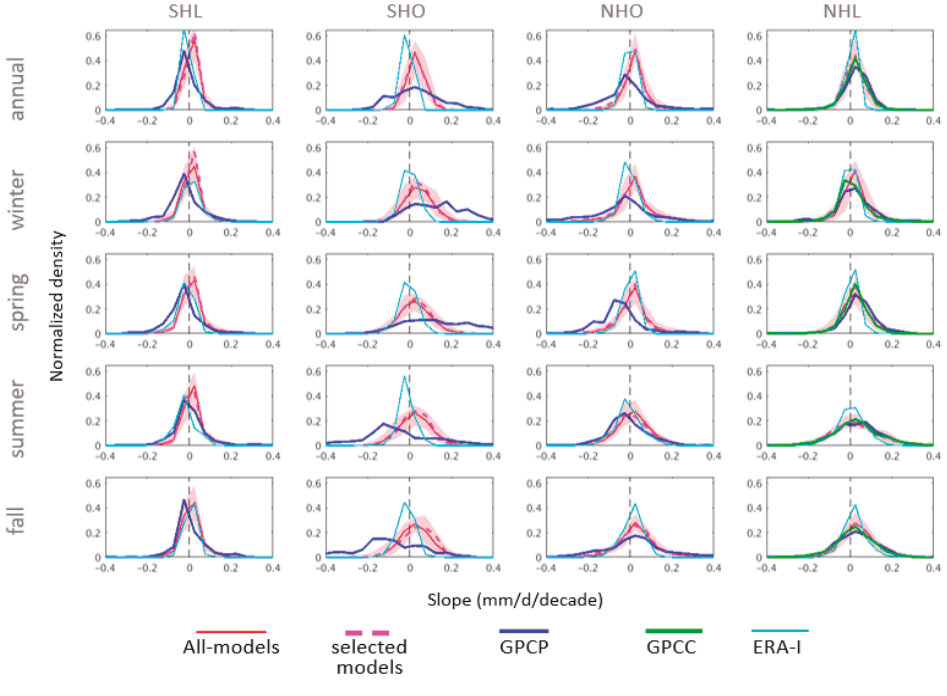


Figure 3. Normalized density plot of precipitation slopes calculated in each grid using 30 years of model (1976–2005), GPCP (1980–2009), GPCC (1980–2009), and ERA-I (1980–2009) data. Models 5–95% ensemble range is shown in pink shades and median as red line. Density plots are shown for four geographical regions. NH land includes land areas north of 55°N and Greenland.

The observed larger rate of precipitation increase over land in cold months is consistent with long-term analysis of station data in NH (e.g., [40]) and with the argument that Clausius–Clapeyron relationship determines the increase of large-scale precipitation in winter, while the availability of moisture is the dominant limiting factor in summer [41,42]. However, the largest fractional increase in precipitation also occurs in the cooler months over oceans where moisture supply is not a strong constraint. This implies stronger warming during winter, such as that occurred in past reanalysis for oceans over 20–90°N [43]. Stronger winter warming, relative to summer warming, is also found across the CMIP5 ensemble for high northern latitudes [2]. This is why a common method for assessing precipitation changes is to consider the percentage change per degree of warming [44], and to distinguish between thermodynamic and dynamic contributions to this change [45].

Our combined RMSE + bias method selects a subset that generally shows more warming and faster winter precipitation increase than that based on bias alone. From this we can conclude that

those models which better represent the current spatial pattern of precipitation result in greater future warming. The field of emergent constraints, in which observable properties of the current climate state are used to infer future changes, may provide insights on this in future. However, we do not speculate here on why because recent research has shown that identifying robust constraints requires intensive investigation and physical backing [46].

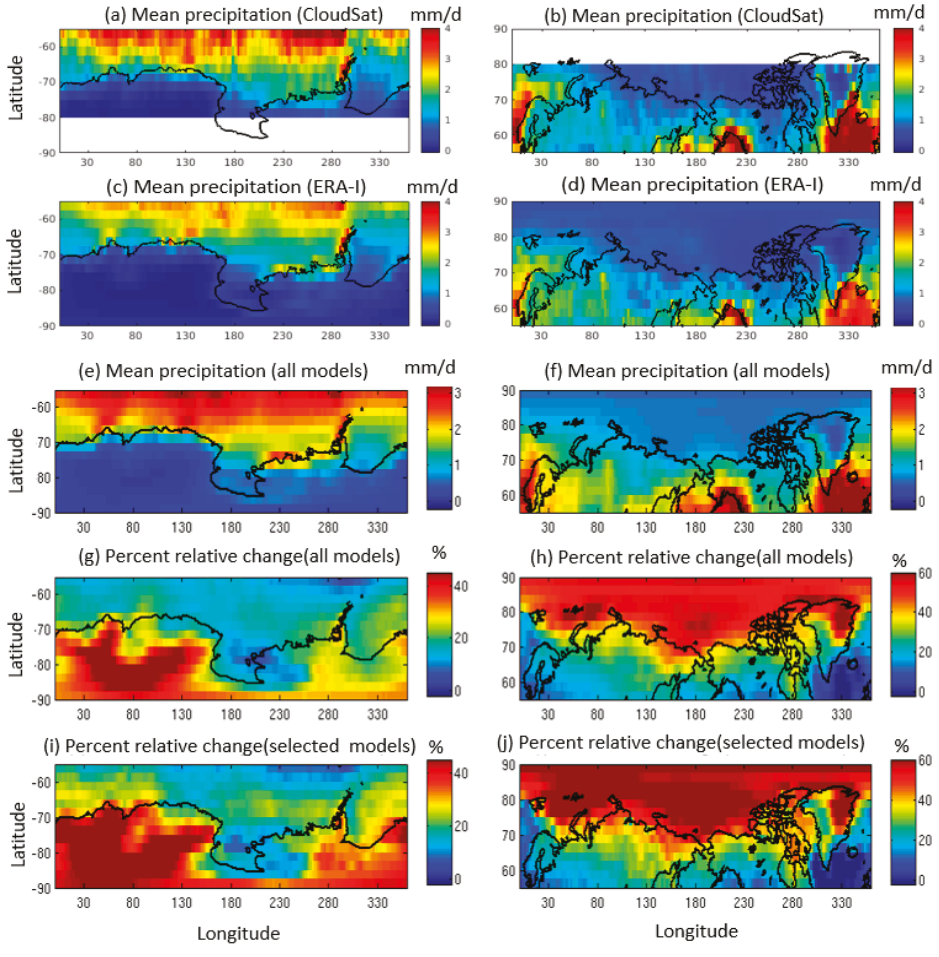


Figure 4. Maps of annual mean precipitation for 2007–2010 from CloudSat (**a**,**b**) and ERA-I (**c**,**d**). It also shows maps of all-model annual mean precipitation for 1976–2005 (**e**,**f**) and percent precipitation change (by comparing 1971–2100 with 1976–2005) for all-model (**g**,**h**) and the selected subset (**i**,**j**) of models in NH and SH. The areas near the North and South poles that are not covered by CloudSat maps are shown in white color.

4.3. Mean Precipitation vs. Surface Temperature Change

The hydrological sensitivity can be defined by the increase in global mean precipitation for a given change in global mean temperature, and this value can be constrained by the global atmospheric energy budget [23]. The local response of this $(\Delta P/P)/\Delta T$ can also be estimated, but is not subject to such constraints due to moisture divergence at the boundaries of the region. However, Figure 6

shows that there tend to be consistent responses across the CMIP5 ensemble for each region. In each case, the subset of selected models is shown by filled circles with the others as empty circles. The full ensemble fits are shown for each region, and for the Antarctic, SH ocean, NH land, NH Ocean, and Greenland are 8.8 ± 1.20, 1.37 ± 1.34, 3.04 ± 1.20, 2.59 ± 0.76, and 7.5 ± 1.46%/°C, respectively. Except for the SH ocean, these are larger than the projected global increase of 2%/°C [47]. This is possible due to moisture transport from other regions and the greater net cooling capacity of the atmosphere at these latitudes relative to the global mean. For example, these high latitudes do not experience super greenhouse effects [48,49] that prevent cooling to space as the local surface warms.

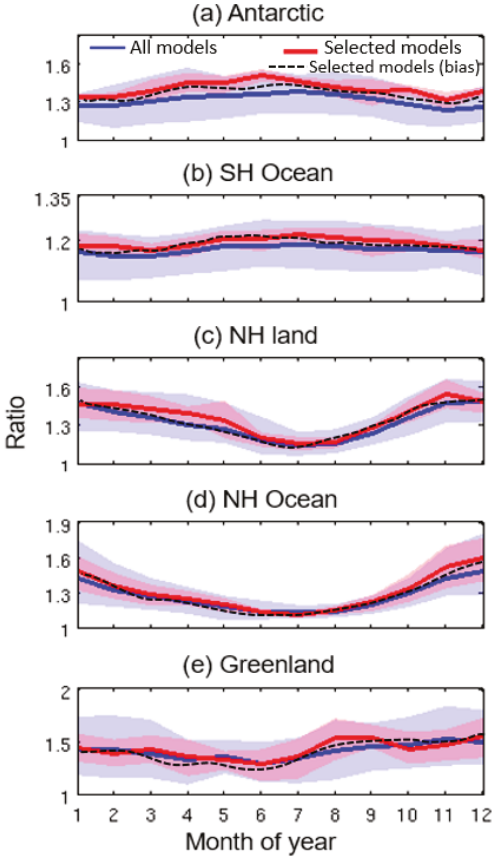


Figure 5. The ratio of future mean precipitation (2071–2100) over the present climate is plotted separately for each month and each region. The median and 20–80% ensemble range of models are shown using line and shaded areas in blue (for all models) and red (for the subset of selected models) in each region.

The largest calculated slope of 8.8 ± 1.20%/°C is over the Antarctic which is higher than the 3–7.4%/°C that has been reported for regional and global climate models (e.g., [24,50]). From the statistics in Table 2, the strongest correlation between ΔP/P and ΔT occurs in the Antarctic (full ensemble r = 0.93) and the weakest in the SH Ocean (full ensemble r = 0.34). This implies a major role for the Clausius–Clapeyron limit on moisture carrying capacity over Antarctica, but limitations based on either moisture availability or changes in horizontal moisture transport through the region

boundaries for the SH oceans. The moisture-availability argument would only apply over sea-ice covered regions.

It can also be seen that the subset of the models generally show larger surface temperature and fractional precipitation change than that full ensemble average, indicating that the larger precipitation changes shown in Figure 6 are largely a result of local thermodynamic effects.

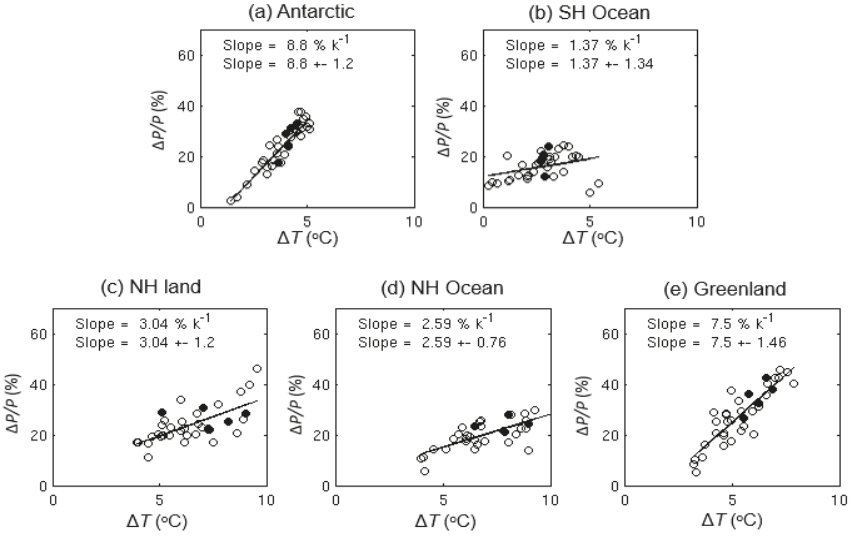


Figure 6. Regional-mean precipitation change (ΔP/P) with respect to surface temperature change (ΔT) for all-models. The subset of selected models in each region is shown by filled-circles, so it can be distinguished from the rest of the models shown by empty circles. The changes are based on comparing the current (1976–2005) and future (1971–2100) climates.

Table 2. Statistics for surface temperature and percent relative precipitation change for all and subset models. The table is constructed using the same data plotted in Figure 6.

		ΔT (°C)			ΔP/P (%)			Correlation
	Models	**min**	**max**	**mean**	**min**	**max**	**mean**	**Coefficient**
Antarctic	All	1.43	5.12	3.79	2.63	37.59	23.26	0.93
	subset	3.68	4.53	4.12	17.46	33.12	26.95	0.89
SH ocean	All	0.27	5.44	2.75	5.68	24.50	15.88	0.34
	subset	2.75	3.06	2.88	12.06	24.07	18.75	0.33
NH land	All	3.21	7.86	5.35	5.44	45.71	27.54	0.87
	subset	5.56	6.92	6.22	26.52	42.68	35.12	0.76
NH ocean	All	3.97	9.58	6.44	11.32	46.29	23.98	0.66
	subset	5.09	9.06	7.34	22.18	30.50	27.12	−0.23
Greenland	All	3.96	10.87	7.12	5.64	32.48	20.64	0.77
	subset	6.48	10.15	8.31	21.29	27.97	24.91	0.53

Figure 7 shows maps of mean near surface temperature for 1976–2005 in SH and NH (Figure 7a,b) and its change through 2071–2100 (Figure 7c,d) for the full ensemble mean. The changes are similar to those shown in Figure 4 for relative precipitation change. While temperature tends to increase almost everywhere, the rate of change has strong regional dependence. The selected subset (Figure 7e,f) has

a similar spatial pattern but often with larger warming, especially over the Arctic Ocean and East Antarctic ice sheet where the greatest fractional changes in precipitation occurred. The NH generally experiences more warming, especially over the historically ice-covered ocean where the increase in mean surface temperature exceeds 10 °C (Figure 7d) in the full ensemble mean or 12 °C (Figure 7f) in the selected model mean.

This fast warming relative to the global mean is consistent with polar amplification arguments [2], with Southern Ocean warming being delayed by factors such as local upwelling of cold ocean waters that reduce warming rates for period of a century or more (e.g., [51,52]).

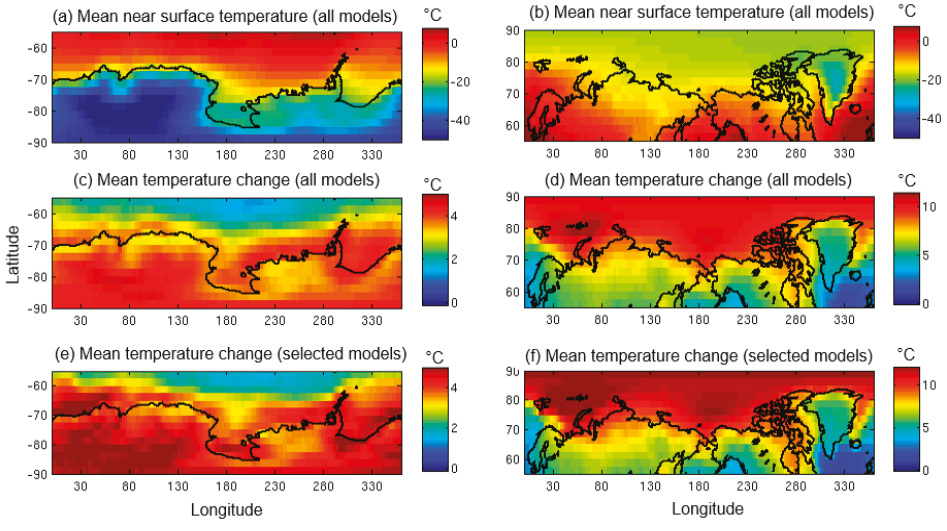


Figure 7. Similar to Figure 4e–j, but for near surface temperature change (°C).

4.4. Changes in Spatial Variability

Figure 8 displays Taylor diagrams with present day in blue linked to future values in magenta by red lines. Arcs of constant RMSE are also displayed, with greater distance from the starred reference dataset indicating greater RMSE. Relative to the bottom left corner, any change that extends the radius represents an increase in the spatial variability of precipitation, and any rotation indicates a change in the spatial pattern. Clockwise means that the future simulated pattern is closer to the present day observed pattern, and counterclockwise the opposite.

Inspection of this figure results in the following interpretations:

(1) Over high latitude oceans, only one model shows present or future spatial variability that is as great as that reported by CloudSat, and that one only over the NH55 ocean region.
(2) The greatest inter-model difference in spatial pattern occurs over high latitude land, since they do not fall on a straight radial line,
(3) Both hemispheres' high latitude land show a general increase in spatial variability. The SH55 land (mainly Antarctica) changes extend radially, indicating little change in the spatial pattern, whereas NH55 land generally show a counterclockwise shift and therefore reduced spatial correlation between present and future.

Point (1) suggests limitations in the simulation of ocean regions of intense precipitation, both in the current and future. Point (2) indicates that the shape and location of the spatial patterns is informative over land. Finally, point (3), with the increased radial extent from present-to-future indicates that these

regions generally experience "wet get wetter", although the angular shift over NH55 land indicates that the location of the wettest and driest regions shifts.

In physical terms, Antarctic precipitation growth mostly occurs on the coast, particularly around the Peninsula and West Antarctic. This occurs in models and in paleoclimate records from historical warmings [50]. The interior is so cold and has such a small current precipitation rate that the faster percentage increase is overwhelmed in terms of mm day^{-1}. The increased spatial standard deviation means that the spread between the wettest and driest regions increases, and with all grid cells in Figure 4 showing increased precipitation this must mean that the wettest areas are getting wetter at a faster rate in absolute terms.

This occurs to some extent over NH land and ocean, but the counterclockwise shift in their future points indicates changes to the pattern of precipitation. This can be inferred from inspection of Figure 4: firstly, the NH ocean south of Greenland shows heavy present day precipitation, but near-zero future changes, for which a number of causes have been proposed [53]. Meanwhile the Bering Sea shows moderate present day precipitation and a substantial future percentage increase. This represents a change in pattern with a decrease in the weighting given to the Atlantic sector. Similar features can be seen over NH land in Figure 4, with major contributions to the correlation between present and future occurring wherever regions of heavy precipitation show small increases, or regions of moderate precipitation show moderate to large fractional increases. Large fractional increases over very dry regions tend to be insufficient to greatly change the spatial correlation. Increases over Finland and the grid cells north of coastal Alaska therefore likely contribute to this change. Further inspection of Figure 4 shows no obvious differences in the spatial patterns, and similar is true for the low-RMSE models in Figure 8. Their future correlation coefficients are not distinguishably different from those of the non-selected models, suggesting that they show similar future patterns of precipitation but with greater mean increase.

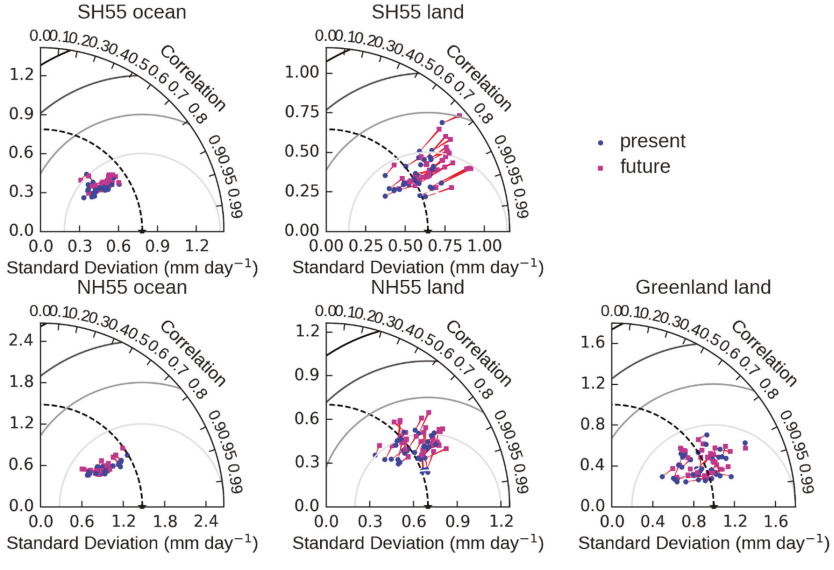


Figure 8. Taylor Diagrams of present and future precipitation for each of the studied high-latitude regions. The reference is shown as a star on the bottom axis in each case, and is CloudSat for SH55 land, SH55 ocean and NH55 ocean. For the others it is ECMWF ERA-I. CMIP5 1976–2005 averages are in blue and 2070–2099 in magenta, with red lines linking each simulation's change.

5. Summary and Concluding Remarks

Behrangi et al. [12] provided a quantitative observation-based update of precipitation amount and distribution poleward of 55° in the Northern and Southern hemispheres (NH55 and SH55) using various data sets including CloudSat over the oceans and Antarctica to take advantage of its high sensitivity W-band radar. Here, this dataset was used to study simulated high-latitude precipitation and assess future changes in precipitation under the RCP8.5 scenario of large global warming. For NH55 land areas, CloudSat's products are inappropriate since rain is frequent there and the rain retrieval relies on path-integrated attenuation, which is currently only determined accurately over ocean surfaces. This forced the use of ECMWF Era Interim reanalysis output over NH55 land, although CloudSat was used over Antarctica since its snowfall product does not use path-integrated attenuation.

We selected a subset of models that we identified as better performing based on the smallest discrepancy relative to CloudSat and ERA-I in terms of both the total amount and spatial distribution of high-latitude precipitation. Models within the five lowest sum of rankings for bias + RMSE relative to the reference for a given region are used in our primary analysis.

The use of minimized RMSE is particularly useful as the CloudSat record is only four years, and internal variability can lead to substantial variation between periods of just four years length, meaning that it can only provide a relatively loose constraint. However, CMIP5 models show greater variation in their calculated regional RMSE than they do in total precipitation. The longer GPCP record was split into non-overlapping 4-year periods to estimate the effect of internal variability on this property, and this was found to be smaller than the typical inter-model differences. This finding suggests that the spatial RMSE allows greater discrimination of models based on the limited time series of available data bias. Meanwhile, we continue to use the additional bias criterion to ensure that the total precipitation is also within a realistic range.

We then consider changes in mean precipitation from 1976–2005 to 2071–2100 for the full CMIP5 ensemble and for a subset of 5 selected based on our bias + RMSE rankings. Any such sub-selection of models can be arbitrary, so we clarify that our sub selection of 5 is illustrative only, and the key result is the extra discriminatory power of using spatial RMSE. When selecting models based on bias alone, the detected changes in precipitation under warming are somewhat smaller than when also considering the spatial pattern. This suggests that analysis including spatial patterns may result in somewhat different constraints than are provided by bias alone.

We showed that while colder months experience larger fractional modelled precipitation increases than warmer months, the selected models generally report larger fractional increase than the full ensemble. For everywhere except the SH ocean, the selected models show greater warming than the model ensemble and tend to fall close to the ensemble local hydrological sensitivity trend, indicating that local thermodynamic effects explain much of the change. For the SH ocean, the models that perform best show temperature changes close to the median but precipitation changes greater than those expected from the full ensemble hydrological relationship, implying that a process other than local thermodynamic changes is the main cause. Based on previous findings [24], this may be related to the sea ice coverage and its change, through the way in which sea ice modifies moisture availability.

A Taylor diagram analysis suggests that across the full CMIP5 ensemble, the NH and SH land areas show increased standard deviation of their spatial precipitation patterns, suggesting a "wet get wetter" paradigm for land poleward of 55°. The SH55 land areas show stable correlations, indicative of small changes in the spatial pattern, but this is not true of NH55 land. This is typical of cases where the spatial pattern of precipitation changes through time as well as the differences in precipitation between wet and dry regions.

Here we presented some potentials and challenges for using CloudSat precipitation estimates to assess the climate models in high latitudes. We note that while CloudSat is more capable than other existing spaceborne sensors in detecting the common type of precipitation in high latitudes (i.e., light rain, drizzle, and snowfall), it may also face various uncertainties, among which is an uncertainty in separating precipition phase. This is partly due to lack of dual-frequency radar as well as the

temperature-based approach that utilizes reanalysis data. While future instruments may reduce such uncertainty sources, the use of wet-bulb temperature [54,55] instead of air temperature might be helpful in the short term. Furthermore, the short period of CloudSat data limits our comprehensive assessment, the near future launch and operation of the Earth, Clouds, Aerosols, and Radiation Explorer mission (EarthCARE) [56] will provide capabilities comparable to CloudSat that helps produce longer data record and enhance our climate model assessment.

Author Contributions: A.B. and M.R. conceived and designed the experiments; A.B. and M.R. analyzed the data and performed the experiment; A.B. and M.R. wrote the paper.

Funding: This research received no external funding.

Acknowledgments: The research described in this paper was carried out at the University of Arizona and Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Financial support was also made available from NASA Energy and Water Cycle Study (NNH13ZDA001N-NEWS), and NASA MEaSUREs (NNH17ZDA001N-MEASURES) awards.

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

References

1. Alexeev, V.A.; Langen, P.L.; Bates, J.R. Polar amplification of surface warming on an aquaplanet in "ghost forcing" experiments without sea ice feedbacks. *Clim. Dyn.* **2005**, *24*, 655–666. [CrossRef]
2. Pithan, F.; Mauritsen, T. Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nat. Geosci.* **2014**, *7*, 181–184. [CrossRef]
3. Min, S.K.; Zhang, X.; Zwiers, F.W.; Agnew, T. Human influence on arctic sea ice detectable from early 1990s onwards. *Geophys. Res. Lett.* **2008**, *35*. [CrossRef]
4. Vaughan, D.G.; Comiso, J.C.; Allison, I.; Carrasco, J.; Kaser, G.; Kwok, R.; Mote, P.; Murray, T.; Paul, F.; Ren, J.; et al. Observations: Cryosphere. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T.F., Ed.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013.
5. Peterson, B.J.; Holmes, R.M.; McClelland, J.W.; Vörösmarty, C.J.; Lammers, R.B.; Shiklomanov, A.I.; Shiklomanov, I.A.; Rahmstorf, S. Increasing river discharge to the arctic ocean. *Science* **2002**, *298*, 2171–2173. [CrossRef] [PubMed]
6. Boening, C.; Lebsock, M.; Landerer, F.; Stephens, G. Snowfall-driven mass change on the east antarctic ice sheet. *Geophys. Res. Lett.* **2012**, *39*. [CrossRef]
7. Zwally, H.J.; Li, J.; Robbins, J.W.; Saba, J.L.; Yi, D.; Brenner, A.C. Mass gains of the antarctic ice sheet exceed losses. *J. Glaciol.* **2017**, *61*, 1019–1036. [CrossRef]
8. Stephens, G.L.; L'Ecuyer, T.; Forbes, R.; Gettelmen, A.; Golaz, J.-C.; Bodas-Salcedo, A.; Suzuki, K.; Gabriel, P.; Haynes, J. Dreary state of precipitation in global models. *J. Geophys. Res. Atmos.* **2010**, *115*. [CrossRef]
9. Adler, R.F.; Gu, G.; Huffman, G.J. Estimating climatological bias errors for the global precipitation climatology project (gpcp). *J. Appl. Meteorol. Climatol.* **2012**, *51*, 84–99. [CrossRef]
10. Berg, W.; L'Ecuyer, T.; Kummerow, C. Rainfall climate regimes: The relationship of regional trmm rainfall biases to the environment. *J. Appl. Meteorol. Climatol.* **2006**, *45*, 434–454. [CrossRef]
11. Lebsock, M.D.; L'Ecuyer, T.S. The retrieval of warm rain from cloudsat. *J. Geophys. Res. Atmos.* **2011**, *116*. [CrossRef]
12. Behrangi, A.; Christensen, M.; Richardson, M.; Lebsock, M.; Stephens, G.; Huffman, G.J.; Bolvin, D.; Adler, R.F.; Gardner, A.; Lambrigtsen, B.; et al. Status of high-latitude precipitation estimates from observations and reanalyses. *J. Geophys. Res. Atmos.* **2016**, *121*, 4468–4486. [CrossRef] [PubMed]
13. Adler, R.; Sapiano, M.; Huffman, G.; Wang, J.-J.; Gu, G.; Bolvin, D.; Chiu, L.; Schneider, U.; Becker, A.; Nelkin, E.; et al. The global precipitation climatology project (GPCP) monthly analysis (new version 2.3) and a review of 2017 global precipitation. *Atmosphere* **2018**, *9*, 138. [CrossRef] [PubMed]
14. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G. Improving the global precipitation record: GPCP version 2.1. *Geophys. Res. Lett.* **2009**, *36*. [CrossRef]

15. Schneider, U.; Becker, A.; Finger, P.; Meyer-Christoffer, A.; Ziese, M.; Rudolf, B. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.* **2013**, *115*, 15–40. [CrossRef]
16. 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*, 1631–1644. [CrossRef]
17. 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]
18. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.K.; et al. Merra: NASA's modern-era retrospective analysis for research and applications. *J. Clim.* **2011**, *24*, 3624–3648. [CrossRef]
19. Bosilovich, M.G.; Robertson, F.R.; Chen, J. Global energy and water budgets in merra. *J. Clim.* **2011**, *24*, 5721–5739. [CrossRef]
20. Kumar, S.; Merwade, V.; Kinter, J.L.; Niyogi, D. Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations. *J. Clim.* **2013**, *26*, 4168–4185. [CrossRef]
21. Gu, G.; Adler, R.F.; Huffman, G.J. Long-term changes/trends in surface temperature and precipitation during the satellite era (1979–2012). *Clim. Dyn.* **2015**, *46*, 1091–1105. [CrossRef]
22. Tapiador, F.J.; Behrangi, A.; Haddad, Z.S.; Katsanos, D.; de Castro, M. Disruptions in precipitation cycles: Attribution to anthropogenic forcing. *J. Geophys. Res. Atmos.* **2016**, *121*, 2161–2177. [CrossRef]
23. DeAngelis, A.M.; Qu, X.; Zelinka, M.D.; Hall, A. An observational radiative constraint on hydrologic cycle intensification. *Nature* **2015**, *528*, 249–253. [CrossRef] [PubMed]
24. Palerme, C.; Genthon, C.; Claud, C.; Kay, J.E.; Wood, N.B.; L'Ecuyer, T. Evaluation of current and projected antarctic precipitation in cmip5 models. *Clim. Dyn.* **2017**, *48*, 225–239. [CrossRef]
25. Hirota, N.; Takayabu, Y.N.; Hamada, A. Reproducibility of summer precipitation over northern Eurasia in CMIP5 multiclimate models. *J. Clim.* **2016**, *29*, 3317–3337. [CrossRef]
26. Behrangi, A.; Lebsock, M.; Wong, S.; Lambrigtsen, B. On the quantification of oceanic rainfall using spaceborne sensors. *J. Geophys. Res. Atmos.* **2012**, *117*. [CrossRef]
27. Behrangi, A.; Tian, Y.; Lambrigtsen, B.H.; Stephens, G.L. What does cloudsat reveal about global land precipitation detection by other spaceborne sensors? *Water Resour. Res.* **2014**, *50*, 4893–4905. [CrossRef]
28. Behrangi, A.; Stephens, G.; Adler, R.F.; Huffman, G.J.; Lambrigtsen, B.; Lebsock, M. An update on the oceanic precipitation rate and its zonal distribution in light of advanced observations from space. *J. Clim.* **2014**, *27*, 3957–3965. [CrossRef]
29. Palerme, C.; Kay, J.E.; Genthon, C.; L'Ecuyer, T.; Wood, N.B.; Claud, C. How much snow falls on the antarctic ice sheet? *Cryosphere Discuss.* **2014**, *8*, 1279–1304. [CrossRef]
30. Haynes, J.M.; L'Ecuyer, T.S.; Stephens, G.L.; Miller, S.D.; Mitrescu, C.; Wood, N.B.; Tanelli, S. Rainfall retrieval over the ocean with spaceborne w-band radar. *J. Geophys. Res.* **2009**, *114*. [CrossRef]
31. Wood, N.B.; L'Ecuyer, T.S.; Heymsfield, A.J.; Stephens, G.L.; Hudak, D.R.; Rodriguez, P. Estimating snow microphysical properties using collocated multisensor observations. *J. Geophys. Res. Atmos.* **2014**, *119*, 8941–8961. [CrossRef]
32. Norin, L.; Devasthale, A.; L'Ecuyer, T.S.; Wood, N.B.; Smalley, M. Intercomparison of snowfall estimates derived from the cloudsat cloud profiling radar and the ground-based weather radar network over Sweden. *Atmos. Meas. Tech.* **2015**, *8*, 5009–5021. [CrossRef]
33. Cao, Q.; Hong, Y.; Chen, S.; Gourley, J.J.; Zhang, J.; Kirstetter, P.E. Snowfall detectability of nasa's cloudsat: The first cross-investigation of its 2c-snow-profile product and national multi-sensor mosaic qpe (nmq) snowfall data. *Prog. Electromagn. Res.* **2014**, *148*, 55–61. [CrossRef]
34. Tanelli, S.; Durden, S.L.; Im, E.; Pak, K.S.; Reinke, D.G.; Partain, P.; Haynes, J.M.; Marchand, R.T. Cloudsat's cloud profiling radar after two years in orbit: Performance, calibration, and processing. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3560–3573. [CrossRef]
35. Grazioli, J.; Madeleine, J.B.; Gallee, H.; Forbes, R.M.; Genthon, C.; Krinner, G.; Berne, A. Katabatic winds diminish precipitation contribution to the antarctic ice mass balance. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10858–10863. [CrossRef] [PubMed]

36. Medley, B.; Joughin, I.; Das, S.B.; Steig, E.J.; Conway, H.; Gogineni, S.; Criscitiello, A.S.; McConnell, J.R.; Smith, B.E.; Broeke, M.R.; et al. Airborne-radar and ice-core observations of annual snow accumulation over thwaites glacier, west antarctica confirm the spatiotemporal variability of global and regional atmospheric models. *Geophys. Res. Lett.* **2013**, *40*, 3649–3654. [CrossRef]
37. Bromwich, D.H.; Nicolas, J.P.; Monaghan, A.J. An assessment of precipitation changes over antarctica and the Southern Ocean since 1989 in contemporary global reanalyses. *J. Clim.* **2011**, *24*, 4189–4209. [CrossRef]
38. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An overview of cmip5 and the experiment design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498. [CrossRef]
39. Moss, R.H.; Edmonds, J.A.; Hibbard, K.A.; Manning, M.R.; Rose, S.K.; van Vuuren, D.P.; Carter, T.R.; Emori, S.; Kainuma, M.; Kram, T.; et al. The next generation of scenarios for climate change research and assessment. *Nature* **2010**, *463*, 747. [CrossRef] [PubMed]
40. Ye, H.; Fetzer, E.J. Atmospheric moisture content associated with surface air temperatures over northern eurasia. *Int. J. Climatol.* **2010**, *30*, 1463–1471. [CrossRef]
41. Berg, P.; Haerter, J.O.; Thejll, P.; Piani, C.; Hagemann, S.; Christensen, J.H. Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]
42. Ye, H.; Fetzer, E.J.; Wong, S.; Behrangi, A.; Olsen, E.T.; Cohen, J.; Lambrigtsen, B.H.; Chen, L. Impact of increased water vapor on precipitation efficiency over northern Eurasia. *Geophys. Res. Lett.* **2014**, *41*, 2941–2947. [CrossRef]
43. Cohen, J.L.; Furtado, J.C.; Barlow, M.; Alexeev, V.A.; Cherry, J.E. Asymmetric seasonal temperature trends. *Geophys. Res. Lett.* **2012**, *39*. [CrossRef]
44. Seager, R.; Naik, N.; Vecchi, G.A. Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming. *J. Clim.* **2010**, *23*, 4651–4668. [CrossRef]
45. Kleidon, A.; Kravitz, B.; Renner, M. The hydrological sensitivity to global warming and solar geoengineering derived from thermodynamic constraints. *Geophys. Res. Lett.* **2015**, *42*, 138–144. [CrossRef]
46. Caldwell, P.M.; Zelinka, M.D.; Klein, S.A. Evaluating emergent constraints on equilibrium climate sensitivity. *J. Clim.* **2018**, *31*, 3921–3942. [CrossRef]
47. Held, I.M.; Soden, B.J. Robust responses of the hydrological cycle to global warming. *J. Clim.* **2006**, *19*, 5686–5699. [CrossRef]
48. Valero, F.P.J.; Collins, W.D.; Pilewskie, P.; Bucholtz, A.; Flatau, P.J. Direct radiometric observations of the water vapor greenhouse effect over the equatorial Pacific Ocean. *Science* **1997**, *275*, 1773–1776. [CrossRef] [PubMed]
49. Stephens, G.L.; Kahn, B.H.; Richardson, M. The super greenhouse effect in a changing climate. *J. Clim.* **2016**, *29*, 5469–5482. [CrossRef]
50. Frieler, K.; Clark, P.U.; He, F.; Buizert, C.; Reese, R.; Ligtenberg, S.R.M.; van den Broeke, M.R.; Winkelmann, R.; Levermann, A. Consistent evidence of increasing antarctic accumulation with warming. *Nat. Clim. Chang.* **2015**, *5*, 348–352. [CrossRef]
51. Armour, K.C.; Bitz, C.M.; Roe, G.H. Time-Varying Climate Sensitivity from Regional Feedbacks. *J. Clim.* **2013**, *26*, 4518–4534. [CrossRef]
52. Rugenstein, M.A.; Caldeira, K.; Knutti, R. Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophys. Res. Lett.* **2016**, *43*, 9877–9885. [CrossRef]
53. Hand, R.; Keenlyside, N.S.; Omrani, N.-E.; Bader, J.; Greatbatch, R.J. The role of local sea surface temperature pattern changes in shaping climate change in the North Atlantic sector. *Clim. Dyn.* **2018**. [CrossRef]
54. Sims, E.M.; Liu, G. A parameterization of the probability of snow-rain transition. *J. Hydrometeorol.* **2015**, *16*, 1466–1477. [CrossRef]

55. Behrangi, A.; Yin, X.; Rajagopal, S.; Stampoulis, D.; Ye, H. On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties, and hydrologic importance. *Q. J. R. Meteorol. Soc.* **2018**. [CrossRef]
56. Illingworth, A.J.; Barker, H.W.; Beljaars, A.; Ceccaldi, M.; Chepfer, H.; Clerbaux, N.; Cole, J.; Delanoë, J.; Domenech, C.; Donovan, D.P.; et al. The earthcare satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 1311–1332. [CrossRef]

remote sensing

MDPI

Article

Cross-Validation of Observations between the GPM Dual-Frequency Precipitation Radar and Ground Based Dual-Polarization Radars

Sounak Kumar Biswas * and V. Chandrasekar

Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA; chandrasekaran.venkatachalam@colostate.edu

* Correspondence: sounak.biswas@colostate.edu; Tel.: +1-970-443-1321

Received: 16 August 2018 ; Accepted: 30 October 2018; Published: 9 November 2018

Abstract: The Global Precipitation Measurement (GPM) mission Core Observatory is equipped with a dual-frequency precipitation radar (DPR) with capability of measuring precipitation simultaneously at frequencies of 13.6 GHz (Ku-band) and 35.5 GHz (Ka-band). Since the GPM-DPR cannot use information from polarization diversity, radar reflectivity factor is the most important parameter used in all retrievals. In this study, GPM's observations of reflectivity at dual-frequency and instantaneous rainfall products are compared quantitatively against dual-polarization ground-based NEXRAD radars from the GPM Validation Network (VN). The ground radars, chosen for this study, are located in the southeastern plains of the U.S.A. with altitudes varying from 5 to 210 m. It is a challenging task to quantitatively compare measurements from space-based and ground-based platforms due to their difference in resolution volumes and viewing geometry. To perform comparisons on a point-to-point basis, radar observations need to be volume matched by averaging data in common volume or by re-sampling data to a common grid system. In this study, a 3-D volume matching technique first proposed by Bolen and Chandrasekar (2003) and later modified by Schwaller and Morris (2011) is applied to both radar data. DPR and ground radar observations and products are cross validated against each other with a large data set. Over 250 GPM overpass cases at 5 NEXRAD locations, starting from April 2014 to June 2018, have been considered. Analysis shows that DPR Ku- and Ka-Band reflectivities are well matched with ground radar with correlation coefficient as high as 0.9 for Ku-band and 0.85 for Ka-band. Ground radar calibration is also checked by observing variation in mean biases of reflectivity between DPR and GR over time. DPR rainfall products are also evaluated. Though DPR underestimates higher rainfall rates in convective cases, its overall performance is found to be satisfactory.

Keywords: GPM; DPR; validation network; volume matching; reflectivity; rainfall rate

1. Introduction

The Tropical Rainfall Measurement Mission (TRMM) satellite carried the first ever space borne weather radar operating at Ku-Band [1]. It was launched in November 1997 as a joint mission by the NASA (National Aeronautics and Space Administration) and the JAXA (Japan Aerospace Exploration Agency). Its main scientific objective was to provide a global 3-D precipitation map over the tropics which would enable the scientific community in extending their research to better understand Earth's energy and hydrological cycle [2]. After a successful span of 17 years, on July 2015, the TRMM satellite came to an end. The Global Precipitation Measurement (GPM) mission is a follow-on mission to the TRMM. Launched on 27 February 2014, the GPM Core Observatory is expected to provide next-generation measurements of precipitation over a wide range of latitudes of 65°N to 65°S [3,4]. The GPM Core Observatory carries a Dual-Frequency Precipitation Radar (DPR), which operates at

13.6 GHz (Ku-Band) and 35.5 GHz (Ka-Band) [5], and a passive microwave radiometer known as GPM Microwave Imager (GMI) [6]. The combined observations of DPR and GMI, extend GPM's capability over TRMM in accurately measuring light precipitation and snowfall [7,8].

Accurate measurement of precipitation is necessary for profound understanding of Earth's hydrometeorological cycle. Satellite measurement of Quantitative Precipitation Estimation (QPE), aided by comprehensive ground validation, can provide reliable and consistent global rainfall data sets over oceans and remote regions [9]. Precipitation measurements from GPM can be very challenging, especially at Ka-band, due to non-Rayleigh scattering effects. Moreover, frequencies at Ku- and Ka-band suffers significant amount of attenuation while propagating through precipitation media. The present GPM precipitation retrieval algorithm [10] is expected to have better performance in quantitatively measuring precipitation over TRMM. It is also expected to improve the estimation of light rainfall by using dual-frequency observations. Even though the retrieval algorithms perform effectively [11] , cross-validation of the GPM measurements and derived products with polarimetric ground radars is of enormous importance for evaluating and improving algorithm performance. It is also crucial for better understanding sources of uncertainties in measurements, consistency in algorithm performance and for quantifying errors. Since GPM pre-launch era, various validation studies have investigated distinct characteristics of GPM-DPR's performance. Kubota et al. [12] found that precipitation estimation using at-launch codes would underestimate Ka-band precipitation in comparison to Ku-band and dual-frequency product. Toyoshima et al. [13] studied storm-top heights measured by the DPR and concluded that the Ka-band radar may not be as sensitive to precipitation measurements as expected. Before launch, dual-frequency algorithms were tested and validated on data simulated at Ka-band from the TRMM-PR's Ku-band measurements [14,15]. Early evaluations showed promising results. In the post-launch period, validation studies of real GPM observations have been conducted with ground radars across the world. Speirs et al. [16] have evaluated GPM-DPR surface rainfall estimate based on different seasons and terrain over Switzerland with C-band radar network. They found the product performs better in flatter terrains and during summer months. In contrast, the performance was significantly degraded over complex terrain, especially in winters. Similar results were also reported by Gabella et al. [17] while evaluating GPM rain-rate products over the Swiss Alps region using ground radar and rain gauge network. Biswas and Chandrasekar [18] performed ground validation of GPM-DPR observations and rainfall rate measurements over the Dallas Fort Worth region in Texas, USA using S-band NEXRAD. They have reported the reflectivities were well matched while rainfall rate product performance was satisfactory. Warren et al. [19] studied ground radar calibration against GPM-DPR Ku-band using 3 operational S-band ground radar near Sydney, Australia. Crisologo et al. [20] also performed studies of calibrating ground radars using GPM over Manila Bay, Philippines and found that DPR-GR bias improved by considering radar beam blockage as a quality filter. Other notable studies have been conducted which deals with validating rainfall estimates from combined GPM-DPR and passive microwave sensor measurements [21–24]. In this work, a large-scale comparison and validation of GPM Ku-band and Ka-band reflectivity measurements and instantaneous rain-rate products has been presented with ground-based polarimetric radars over the United States.

NASA has developed a Ground Validation (GV) program for doing scale large validation in the United States [25]. At present, the NASA Validation Network (VN) consists of 75 operational NEXRAD sites [26]. This study uses 5 NEXRAD radars located in the southeastern plains of the U.S.A. NEXRAD radars are chosen since they operate at a non-attenuating frequency of S-band. At the same time, GPM undergoes severe attenuation making is difficult for direct comparison. The GPM reflectivity measurements first need to be corrected for attenuation. This is achieved by the algorithm in [10,27]. To minimize errors from temporal mismatch both GR and GPM overpass data are carefully chosen such that the relative time difference is less than 5 min. Both data are volume matched following the procedure mentioned in [25,28]. GPM-DPR reflectivity factor at Ku- and Ka-band along with instantaneous rain-rate product are cross validated using 4-year (April 2014 to June 2018) dataset

of GPM overpasses with the mentioned GV sites. Radar reflectivity factor is chosen because it is the fundamental product used to retrieve precipitation from radar. In absence of any attenuation, an ideal relationship between GR S-band, GPM-DPR Ku- and Ka-band frequencies was determined by performing scattering simulations. Results from simulation are used to measure the theoretical deviation in GPM-DPR reflectivity for a given mean GR reflectivity. This in turn is used to access of DPR-GR biases. The main aim of this study is to quantify the difference in reflectivity and to access the performance of instantaneous rain-rate product at a large scale over the United States using GV radars from NASA ground validation network. This article is organized as follows. In Section 2 the data sets from GPM and NEXRAD radars are summarized. Volume matching methodology and the scattering simulations are also presented. Section 3 describes quantitative comparisons of the GPM-DPR measurements with respect to ground radars. Absolute calibration of ground radars against GPM is also studied. Lastly, Section 4 provides a summary of this work along with discussions and conclusions.

2. Instruments, Datasets and Analysis Methodology

2.1. The GPM Satellite: Dual-Frequency Radar

The heart of the GPM core observatory is the DPR and the GMI. The combined radar-radiometer observations are expected to improve precipitation estimates. The DPR consists of two radars namely KuPR operating at 13.6 GHz and KaPR operating at 35.5 GHz. [5] They are capable of providing high-resolution three-dimensional precipitation data. Figure 1a depicts DPR's observational characteristics. It can be observed that below melting layer significant difference in attenuation suffered at the two frequency channels. One of the main reasons for adding the Ka-band channel is to improve the detection threshold for light rain and snow compared to TRMM's precipitation radar [3,7]. Non-Rayleigh scattering effects are more pronounced at Ka-band than in Ku-band. This phenomenon coupled with the difference in attenuation makes a feasible parameter for detailed study of precipitation microphysics. Using this dual-frequency information enables improved drop size distribution retrievals [29], accurate melting layer height detection [30] and identification of different hydrometeor phases such as snow [8,31].

Each radar consists of active phased array elements (128 slot array antennas) capable of 49 simultaneous beam formations. The antenna and the pulse repetition time is designed such that the DPR can observe at least 19 Km from the sea level. Figure 1b shows the scan strategy. The KuPR radar scans with all the 49 beams and it known as normal scan (NS). Each footprint has a diameter of approximately 4.9 Km at NADIR which gives a scan swath of 245 Km. The KaPR radar scan is matched across the central 125 Km of the KuPR swath. This scan is called the Matched Scan (MS). The vertical range resolution of both the radars is 250 m. There is another type of scan called the High Sensitivity Scan (HS) where all the 49 beams of the KaPR radar are interlaced with the middle 25 beams of the KuPR radar. During signal processing, the radar echoes are over sampled at twice the rate, giving a vertical resolution of 125 m in the final data. The narrow bandwidth of the KaPR can be accounted for side lobe clutter contamination for large scan angles. A detailed design specification can be found in the work of Senbokuya et al. [5].

The GPM data used in this study is 2ADPR level 2 data, version 5 [32]. It should be noted that there has been changes in DPR calibration, from previous version 4 level 2 data, corresponding to increase in KuPR reflectivity by 1.2 dB and KaPR reflectivity by 1.3 dB [33]. Over 250 DPR overpass data that have been used in this study were downloaded from the STORM online data interface to NASA Precipitation Processing System archive (https://storm.pps.eosdis.nasa.gov). Products from only Normal and MSs are used. Highly sensitive scan data is not considered to avoid complications arising from the geometry of interlaced beams. The key variables used in this work include zFactorMeasured from Preparation Module, zFactorCorrected, precipRate and precipRateNearSurface from Solver Module and typePrecip, qualityTypePrecip, heightBB and flagBB from Classification Module.

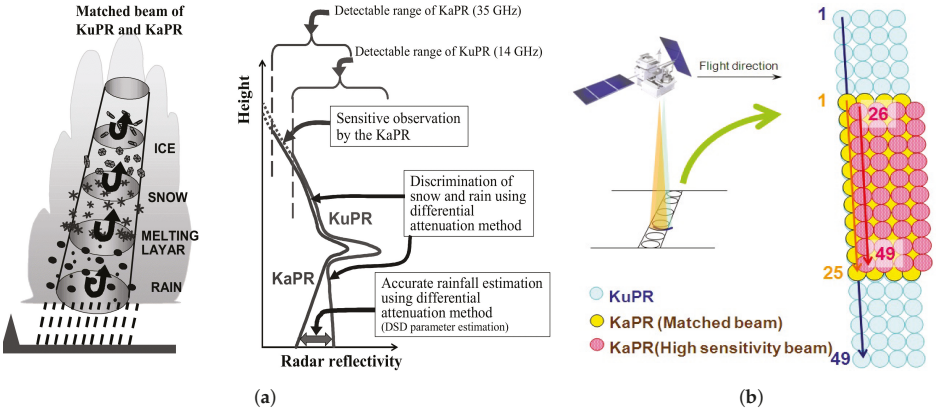


Figure 1. (**a**) Concept of DPR dual-frequnecy observation. Image source: Senbokuya et al. [5]; (**b**) GPM-DPR scan strategy. Image taken from GPM Level 2 algorithm theoretical basis document [10].

2.2. Ground Validation Radar: NEXRAD

The Weather Surveillance Radar, 1988, Doppler or WSR-88D are S-band operational radars used for weather observation and forecasting in the United States. They are also known as the Next-Generation Weather radar or NEXRAD. These radars are operated by National Oceanic and Atmospheric Administration (NOAA) and Department of Defense (DoD). At present there is a deployment of 160 WSR-88D, throughout the continental USA. Recently, NEXRAD radar have been upgraded with the capability of polarization diversity. They operate at the frequency range of 2700 MHz to 3000 MHz. The radar has a range resolution of 250 m. The beam is 1 degree wide. Typical range of NEXRADs is 400 Km. More information on NEXRAD specifications can be found in [34]. During normal operations, the radar conducts a volume coverage pattern (VCP) in which it scans full 360 degree in azimuth and from 0.5 degree to 19.5 degree in elevation in steps of 14 plan position indicator (PPI) scans. There are different variety of VCP which are optimized for specific weather conditions and are deployed by the operational controls according to the observed meteorological event and desired data quality. The radar typically takes around 5 to 6 min. to complete one volume scan. In complex topography, radar observations are associated with several issues such as partial or total beam blockage and beam overshooting. Current operational radars in the United States do not provide accurate measurement of rainfall over mountainous region. [35] Due to this reason NEXRAD radars located in the southeastern plains along the Gulf of Mexico are chosen. Table 1 provides the names and locations of the 5 NEXRAD radars used for validation. All the radars are located from 5 to 210 m in altitude from Mean Sea Level (MSL). NEXRAD level 2 data has been used in this study. Differential reflectivity (Z_{DR}) calibration was carried out from observations of Z and Z_{DR} in drizzle following the procedure mentioned in [36]. It is important to ensure accurate rainfall estimation by dual-pol estimators such as $R(Z, Z_{DR})$ and $R(K_{DP}, Z_{DR})$. Attenuation correction is not necessary since S-band does not suffer attenuation in precipitation medium.

Table 1. NEXRAD radar stations' name, location, and altitude (MSL).

Name	Location	Lat (deg.)	Lon (deg.)	Altitude (m)
KFWS	Dallas/Ft. Worth, TX	32.5731	−97.3031	208.17
KHGX	Houston/Galveston, TX	29.4719	−95.0792	5.48
KSHV	Shreveport, LA	32.4508	−93.8414	83.21
KLIX	New Orleans, LA	30.3367	−89.8256	7.31
KMLB	Melbourne, FL	28.1133	−28.6542	10.66

2.3. Theoretical Simulations of Reflectivities at S, Ku, and Ka-Band

In Rayleigh scattering, the particle sizes are much smaller compared to the radar wavelength. In this case, the radar reflectivity factor can be shown approximately equivalent to the sixth moment of the drop size distribution. This assumption is valid for NEXRAD radars operating at a wavelength of about 10 cm. When particle sizes are greater than 10 percent of the radar wavelength, this Rayleigh approximation does not hold true. If raindrop dimensions are compared against GPM-DPR Ku-band (13.6 GHz) and Ka-band (35.5 GHz), non-Rayleigh scattering effects seem to be very evident in GPM observations at these two frequencies. Due to this, even in ideal conditions (with no attenuation), the back scattering cross section from precipitation particles can be quite different at both GPM Ku-band and Ka-band compared to ground radar S-band. This in turn may lead to inconsistency while comparing the reflectivities. Hence to correctly infer the differences in reflectivities, it is useful perform theoretical simulations. In this study, scattering simulations are done using T-Matrix method [37]. It was assumed that the drops were symmetric with orientation canting angle following a Gaussian distribution with zero mean and standard deviation of 5°. The drop shape model described in [38] has been used. Mean reflectivities were calculated varying the ranges of normalized intercept parameter, shape parameter and median volume diameter of a normalized gamma distribution model [39]. Equation (1) shows a normalized gamma distribution and Equation (2) shows the ranges of N_w, μ and D_o used in the simulation.

$$N(D) = N_w f(\mu) \left(\frac{D}{D_o}\right)^{\mu} e^{-\Lambda D} \tag{1}$$

$$\Lambda = \frac{3.67+\mu}{D_o}$$

$$f(\mu) = \frac{6}{3.67^4} \frac{(3.67+\mu)^{\mu+4}}{\Gamma(\mu+4)}$$

$$\begin{aligned} 10^3 < N_w < 10^5 \\ -1 < \mu < 3 \\ 0.5 < D_o < 2.5mm \end{aligned} \tag{2}$$

The difference between S, Ku, and Ka-band reflectivities obtained from the simulation is shown in Figure 2. It can be observed that Ku-band reflectivity values up to 35 dBZ are almost equal to S-band while Ka-band reflectivity values are almost 1 dB higher than S-band. From 35 dBZ onward, due to increase in drop size, non-Rayleigh scattering effect start to crop in and thus a deviation can be seen in both Ku-band and Ka-band. Beyond 50 dBZ this deviation becomes more severe. In the region of 35–50 dBZ, Ku-band reflectivity can be 2 to 3 dBZ higher while Ka-band reflectivities can be 5 dBZ lower compared to S-band. These results are consistent with the findings in [40,41].

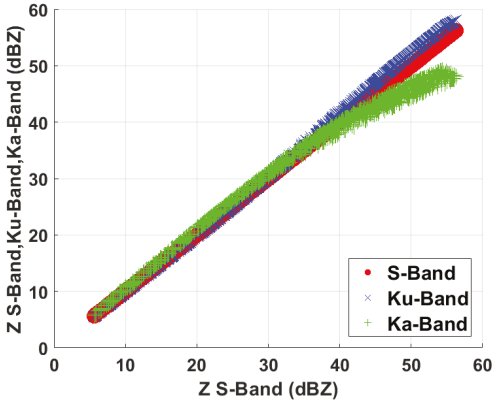


Figure 2. Relationship between reflectivities simulated at S, Ku, and Ka-band for raindrops using T-Matrix method.

2.4. Volume Matching Methodology

Direct inter-comparisons between the GPM-DPR and ground radar is not possible because of the different viewing geometry of the two systems. The space-based radar is a vertically down looking radar while ground-based radars scan in PPI mode. For point by point comparisons, errors can crop in from observation resolution volume mismatch, spatial alignment, and difference in operating frequency.

Figure 3a,b both show the general concept of DPR and GR observation volumes at the beam intersection locations. It can be observed in Figure 3b that these common volumes at different beam intersections can be different. In literature, several procedures have been proposed for matching and aligning space and ground radar observations [42,43]. In the method, the data from space radar and ground radar are interpolated and re-sampled to a common Cartesian grid. A similar methodology by Bolen and Chandrasekar [28,41] discusses the distortion of reflectivity map from space radar with respect to ground radar and its mitigation by re-sampling them to a common grid by variable volume matching with a polynomial technique for alignment. This method was further developed by Schwaller and Morris [25] where volume matching is performed at each geometric intersection of the space radar and ground radar beams by averaging the reflectivity samples within the volume instead of re-sampling it. It has been shown by the authors that this procedure minimizes error due to interpolation. In this study, the same procedure has been followed. First the intersection locations of DPR's beam with GR's elevation sweeps are identified. Only the intersection samples lying within the 100 Km range radius of the ground radar are considered. This is because, at ranges beyond 100 Km, the vertical extend of the GR beam is nearly 2 Km which is considered too coarse for meaningful comparisons. Moreover, due to nonstandard atmospheric refraction the GR beam bends. The vertical resolution of the DPR data is 125 m and the horizontal resolution is 4.9 Km whereas the resolution of GR data is 250m. Thus, DPR gates falling between the half-power GR beam points are linearly averaged in vertical direction. The GR data are averaged in horizontal direction over a circular area centering the intersection of parallax adjusted DPR beam with GR sweeps. The GR gates are weighted with a Barnes [44] type Gaussian inverse distance weighting. Here the distance is measured horizontally from the center of DPR footprint to GR gates' center. This is necessary to account for the non-uniform power distribution within the DPR beam. Due to this process, vertical resolution of DPR data and horizontal resolution of GR data are both decreased. This procedure produces an output which is a set of vertical profiles aligned along the DPR beam with samples points located at height of each intersection of GR sweep with DPR beam. All variables from GR and DPR are averaged following the

same procedure. The reflectivity factors are averaged in linear scale and converted back to logarithmic scale. The minimum detection threshold for both KuPR and KaPR is 18 dBZ [3]. Therefore, reflectivity values above this level are only considered in averaging within the matched volume.

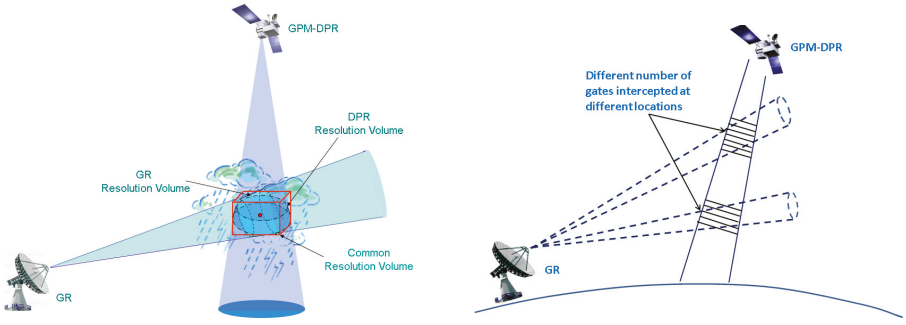


Figure 3. Diagram illustrating viewing geometry of DPR and GR. (**a**) Intersection of DPR and GR common volumes. (**b**) Cartoon showing different number of DPR gates at common intersection locations for different GR sweeps. Image adapted from [25].

3. Results and discussions

Over a total of 250 GPM overpass cases with NEXRAD radars mentioned in Section 2.2 have been studied. Precipitation data from April 2014 to June 2018 have been used. Reflectivities without attenuation correction from DPR are not considered in the comparison. GPM overpasses are carefully chosen such that enough precipitation was present within the intersection of KuPR swath and 100 Km range of GR. Three statistical indices are selected for evaluating GPM observations and products with respect to GR. Mean Bias (MB) is used to observe systematic bias in measurements. Mean Absolute Error (MAE) is used to quantify the average magnitude of error. Pearson's correlation coefficient is used to measure how well the two observations match with each other. Mathematically they can be written as

$$MB = \frac{1}{N}\sum_{n=1}^{N}(Z_{DPRn} - Z_{GRn}) \tag{3}$$

$$MAE = \frac{1}{N}\left|\sum_{n=1}^{N}(Z_{DPR} - Z_{GR})\right| \tag{4}$$

$$Corr = \frac{\sum_{n=1}^{N}(Z_{GR} - \overline{Z_{GR}})(Z_{DPR} - \overline{Z_{DPR}})}{\sqrt{\sum_{n=1}^{N}(Z_{GR} - \overline{Z_{GR}})^2(Z_{DPR} - \overline{Z_{DPR}})^2}} \tag{5}$$

where N is the total number of samples and the subscript n implies nth sample. Since Z_{GR} and Z_{DPR} are in logarithm scale, $\overline{Z_{GR}} = \frac{1}{N}\sum_{n=1}^{N} Z_{GRn}$ is actually the geometric mean of equivalent reflectivity factor from all GR samples. Similarly, $\overline{Z_{DPR}} = \frac{1}{N}\sum_{n=1}^{N} Z_{DPRn}$ is the geometric mean of equivalent reflectivity factor from all samples from DPR.

3.1. Comparison of Radar Reflectivity

Volume matched reflectivity from NEXRAD radars is compared against volume matched attenuation corrected reflectivity from DPR Ku-band and Ka-band band. Volume matched samples

below bright band are chosen so that the comparison is limited to rain region. Figures 4 and 5 shows example of DPR and GR reflectivity scatter diagrams at the 5 NEXRAD location sites for all GPM overpass data samples.

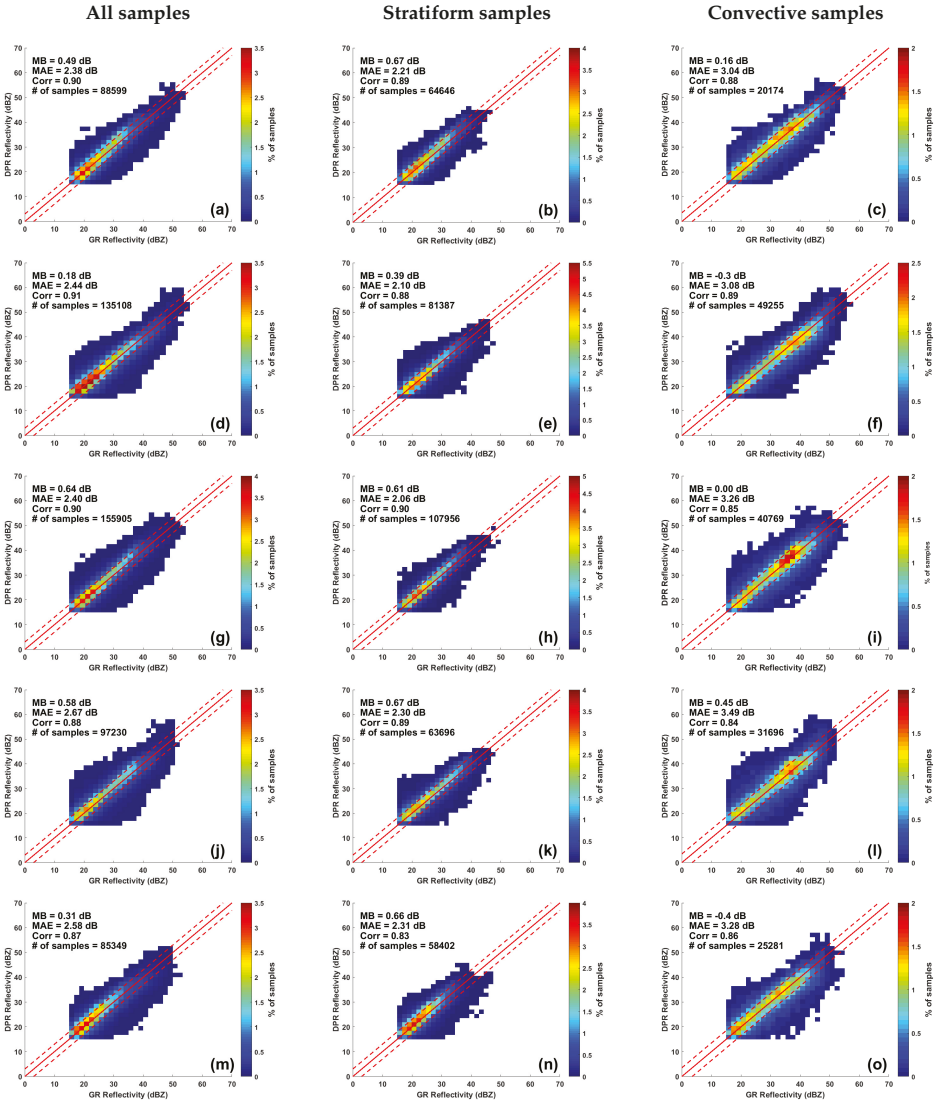


Figure 4. Scatter density plots of volume matched reflectivity between GPM-DPR Ku-Band and NEXRADs. Solid red line is 1:1 line. Dotted red lines are ±3 dB lines. (**a**–**c**) KFWS radar; (**d**–**f**) KHGX radar; (**g**–**i**) KSHV radar; (**j**–**l**) KLIX radar; (**m**–**o**) KMLB radar.

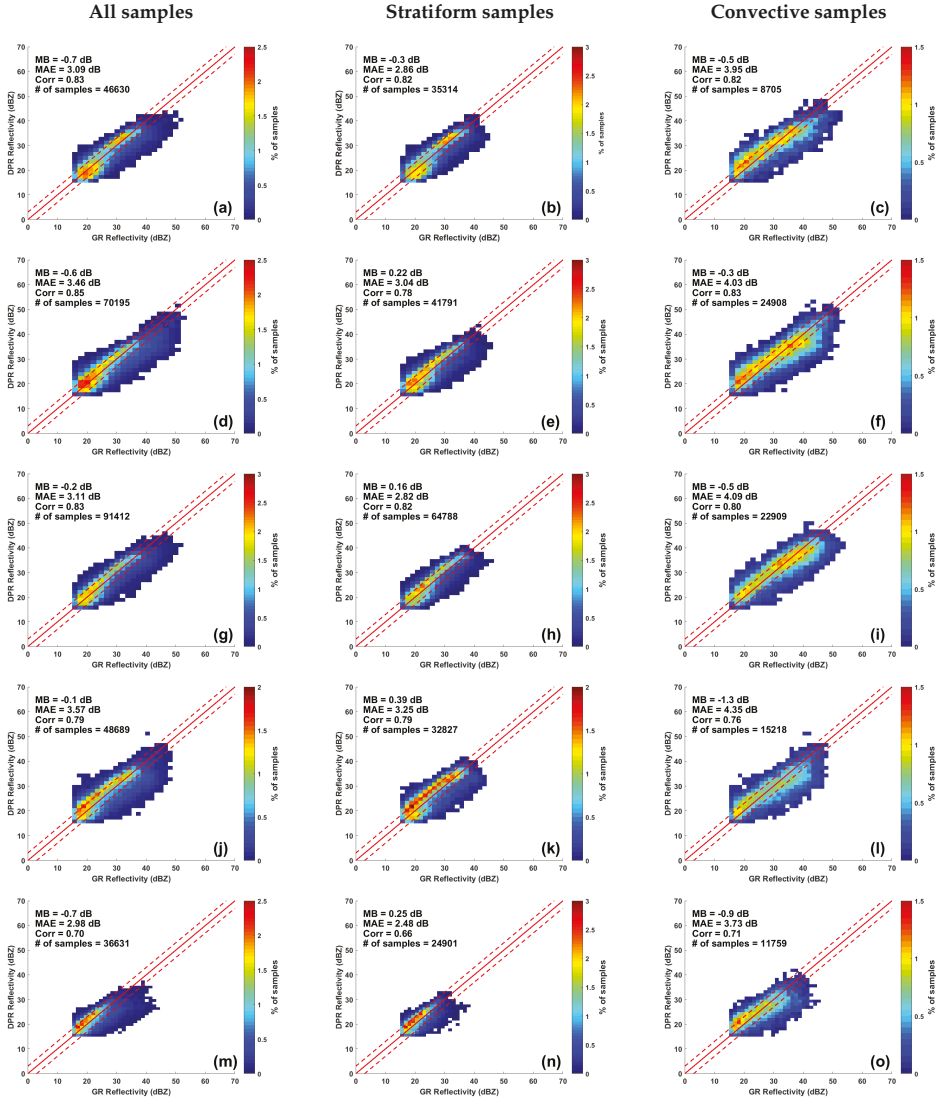


Figure 5. Same as Figure 4 but for GPM-DPR Ka-Band.

Attenuation in DPR outer swath is corrected by a combination of Surface Reference Technique (SRT) and Hitschfeld and Bordan (HB) method [45] while in inner swath a modified HB-SRT-DFR method [27] is used. These correction methods do not always work well for heavy precipitation [27]. Thus, to better understand the effectiveness of attenuation correction, volume matched samples are further classified into stratiform and convective rain. DPR data are flagged according to the precipType flag from Profile Classification Module of the DPR Level 2 algorithm. A dual-frequency method [46] is applied in the inner swath where both Ku- and Ka-band data are available. In the outer swath, a single frequency approach [47] is adopted.For GR data classification, matched volume samples are chosen according to the DPR classification algorithm. The first column of both Figure 4 shows scatter density plot between attenuation corrected KuPR reflectivity and GR reflectivity when all volume matched

samples are considered together. The second and third column are similar but only for stratiform and convective volume matched samples, respectively. Figure 5 is same as Figure 4 but for attenuation corrected KaPR reflectivities. Statistical scores together with the number of samples used are shown on each plot. From both Figures 4 and 5, It can be observed that, even though there are differences between DPR and GR matched samples, data are heavily clustered along the 1:1 line and within the bounds of ±3 dB lines. Statistical scores shown on the scatter plots are summarized in Table 2.

Table 2. Table summarizing different error statistics for comparison of DPR Ku-band and Ka-band reflectivity vs NEXRAD S-band reflectivity.

	All Samples						Stratiform Samples						Convective Samples					
	Ku Band (dBZ)			Ka Band (dBZ)			Ku Band (dBZ)			Ka Band (dBZ)			Ku Band (dBZ)			Ka Band (dBZ)		
Radar	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**
KFWS	0.49	2.38	0.90	−0.7	3.09	0.83	0.67	2.21	0.89	−0.3	2.86	0.82	0.16	3.04	0.88	−0.5	3.95	0.82
KHGX	0.18	2.44	0.91	−0.6	3.46	0.85	0.39	2.10	0.88	0.22	3.04	0.78	−0.3	3.08	0.89	−0.3	4.03	0.83
KSHV	0.64	2.40	0.90	−0.2	3.11	0.83	0.61	2.06	0.90	0.16	2.82	0.82	0.0	3.26	0.85	−0.5	4.09	0.80
KLIX	0.58	2.67	0.88	−0.1	3.57	0.79	0.67	2.30	0.89	0.39	3.25	0.79	0.45	3.49	0.84	−1.3	4.35	0.76
KMLB	0.31	2.58	0.87	−0.7	2.98	0.70	0.66	2.31	0.83	0.25	2.48	0.66	−0.4	3.28	0.86	−0.9	3.73	0.71

In Section 2.3, it is shown that Ku-band reflectivities can be 2 to 3 dB greater in rain region. From Table 2, it can be seen, when all samples are considered, for GPM Ku-band an overall positive MB near to zero is observed at each GR site. The MAE lies between 2 and 3 dB. These systematic differences agree with the simulation results. For all radars, the correlation between samples lie around 0.9 which indicate the DPR Ku-band reflectivities are well matched with GR. Similar results for MB and correlation are also reported in [19,20]. When Ka-band is considered, a negative MB can be seen for all radar sites, which means that the overall reflectivities observed at Ka-band is lower than GR S-band. This is also is agreement to the simulation study presented in Section 2.3. The correlation between matched samples at all the radar sites lie between 0.7 and 0.85 implying that with respect to GR, Ka-band samples are not matched as good as Ku-band. The decrease in correlation can be accounted for the attenuation correction algorithm not being as effective as it is for Ku-band. Lower correlations also affected the MAE for Ka-band which are larger compared to Ku-band and lies in the range of 3 to 3.5 dB. However, the errors are still within bounds for the Ka-band and S-band reflectivity difference reported in Section 2.3. When rain profile classification is considered, the correlations between DPR Ku-band and GR matched stratiform samples at all radar sites is observed to be almost similar to that of convective samples. However, the mean correlation considering all radar sites is found to differ by a small margin of 0.03. Similar trend can be observed at Ka-band. The mean biases of Ku-band stratiform samples at all radar sites is positive whereas, two deviations can be observed for Ku-band convective samples for radars KHGX and KMLB. Both these radars are located at sea level and experiences heavy convective events resulting from tropical storms. The attenuation correction may be underestimating the true reflectivity giving rise to negative biases. It is interesting to note that the mean biases of convective samples at Ka-band are negative whereas they show a positive trend in case of stratiform samples other than at KFWS radar site. In case of stratiform rain, reflectivity at S-band is generally lower than 35 dBZ. The scattering simulation presented in Section 2.3 shows that Ka-Band reflectivity can be up to 1 dBZ higher corresponding to S-band range of 15 dBZ to 35 dBZ. This can be correlated with the positive trend. The MAE of both DPR Ku-band and Ka-band for both types of rain are within the difference limit as noted from scattering simulation results. It is also worthwhile to mention that the GPM profile classification algorithm also classifies rain into a third category named "other" for profiles which are neither stratiform nor convective [46]. In this study, the 'all samples' case contains samples from these 3 types of rainfall. Case study for 'other' type classification is not shown separately since this type contributes a very small percentage to the total rainfall profiles. However, nevertheless, when considered together with the other two types, correlation is improved by a very small factor in most cases.

3.2. Absolute Calibration of Ground Radars with Respect to DPR

One of the potential useful applications of the GPM-DPR data is to track the absolute calibration of ground radars over time. In this study, calibration errors of ground radars are studied with respect to the DPR Ku-band measurements. Figure 6 presents the DPR-GR mean difference for each month for all 5 NEXRAD radars over a period of year 2014 to 2018. When considering bias month wise, the Ka-band sample size can be very low in each case. It is very difficult to have GPM overpasses which has significant amount of precipitation present in the inner swath. Moreover, in many cases, the inner swath falls outside the 100 Km range of GR. Due to these reasons along with the performance issues of attenuation correction at Ka-band, only Ku-band has been considered for studying the time series of monthly bias.

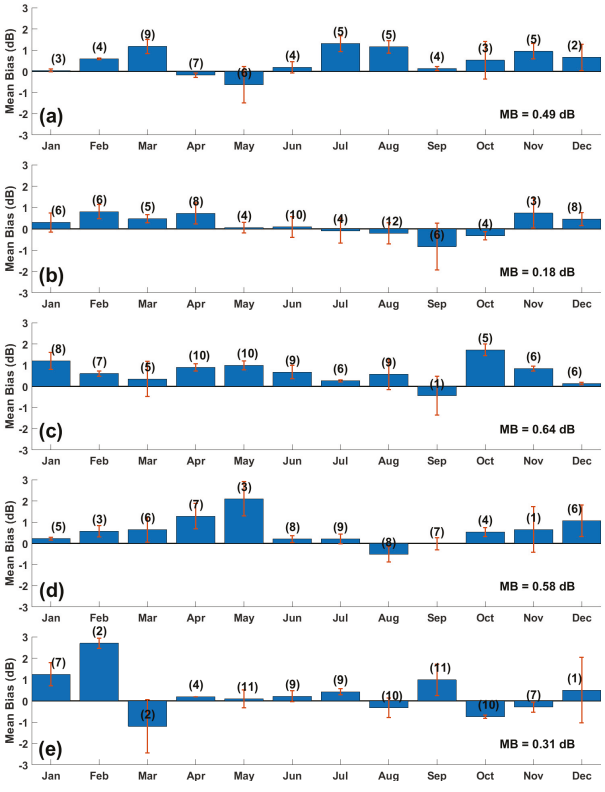


Figure 6. Bar diagram of monthly DPR-GR mean reflectivity bias with standard deviation for the years 2014 to 2018. Numbers in round brackets signify total number of overpasses for that month from April 2014 to June 2018. (**a**) KFWS radar; (**b**) KHGX radar; (**c**) KSHV radar; (**d**) KLIX radar; (**e**) KMLB radar.

In Figure 6, MB for each radar over the months are shown in blue bars with red lines on top representing standard deviation. It can be seen that the calibration of all radars is relatively stable over the time period with occasional jumps of a few dB at some radar sites. The MB is seen to vary from −1.1 dB to 2.5 dB. Other than factors such as difference in back scattered power at higher frequency and DPR attenuation correction algorithm performance, rainfall climatology can be a significant factor for the variability of MB over different months. The southeastern part of USA experiences a wide range of weather events, from severe thunderstorms, wide spread stratiform and mesoscale convective systems to tropical cyclones. During spring, sharp changes in temperature gradient along with varying

humidity facilitate development of hail producing thunderstorms. The summer season is typically warm and moist and characterized by frequent convective thundershowers. During late summer, tropical cyclones also cause torrential rainfall upon interacting with frontal systems [48]. During the cooler months, it experiences comparatively light rainfall from widespread stratiform events. The DPR-GR bias is mostly positive and near to zero during this season which can be noticed in case of KFWS, KSHV and KLIX radars. The KMLB radar shows an exception of high positive bias for the month of February. It should be noted that there are only two overpasses for this month which might not be sufficient for truly capturing the nature of the bias. Also, in convective precipitation, when the Ku-band signal attenuation is significant the correction algorithm can sometime change the raw radar reflectivity by a factor of 10 compared to the corrected value (Iguchi et al. 2000). During light rain events, the DPR's SRT can sometimes yield negative amount of attenuation correction (Schwallar and Morris 2011). This might be a reason of negative DPR-GR bias which can be noticed for some of the cool season months for radars KHGX, KLIX and KMLB. A noticeable deviation can be observed in case of KFWS radar which shows negative bias for the month of May. The Dallas Fort Worth region is known for observing damaging hail storms during the summer. This can be the cause for the heavy attenuation of Ku-band signal and poor performance of the attenuation correction algorithm. Nevertheless, errors in GR hardware systems during these overpass events can be another cause for sudden jumps in biases.

3.3. Comparison of Instantaneous Rainfall Rate

Volume matched samples of instantaneous rainfall rate product from GPM-DPR are compared against volume matched samples of retrieved rainfall rate from ground radar. Rainfall rate retrieval from DPR is based on the algorithm described in [10,11]. In the outer swath Ku-only retrieval process [10] is used similar to TRMM whereas in the inner swath a dual-frequency approach [10,29] is adopted. The algorithm used to retrieve rainfall from ground radars is based on [49] which is a blended product of different dual-pol estimators. More detailed description of the algorithm can be found in [49]. This algorithm has been developed for precipitation characteristic in the southern and southeastern plains. This algorithm's high performance with respect to rain gauges is also shown in [49].

Figure 7 shows the scatter density plots between matched samples of DPR Ku-only rain-rate and GR rain-rate at each GR site. Figure 8 also shows scatter density plots of matched rain-rates samples but with DPR dual-frequency product. The samples are again classified according to rain type. The first column of both Figures 7 and 8 represents scatter plots from all matched samples, while the second column and third column represent stratiform and convective samples, respectively. It can be seen that lower rain-rate samples (0.1 to 3 mm/h) are more clustered around the 1:1 line. For higher rain-rates the samples diverge from 1:1 line. Statistical induces are summarized in Table 3. Mean biases for the Ku-only product varies between -1.7 to 0.3 mm/h. The corresponding mean biases for the dual-frequency product decreased for all the radar sites except KSHV. Overall, this implies that both DPR Ku-only and dual-frequency underestimates rainfall expect at KSHV site. The MAE for Ku-only product varies between 1.66 to 2.56 mm/h whereas for dual-frequency product the error magnitude increases which is from 1.69 to 2.98 mm/h. The correlation of the matched samples for both products varies in the range of 0.5 to 0.7. This result, particularly for Ku-only product, is consistent with findings from [50–52]. When considering stratiform samples alone, mean biases are found to decrease. They are in the range of 0.13 to 0.43 mm/h for Ku-only product and 0.02 to 0.81 mm/h for dual-frequency products. The MAEs are also lower and are in range of 1 to 1.5 mm/h. However, for convective cases both products do not show satisfactory performance compared to GR rain-rate estimates. The error magnitude is as high as 5.56 mm/h for Ku-only product and 6.10 mm/h for dual-frequency product. Except at the KSHV radar site, mean biases for both products are also seen to increase towards negative indicating high underestimation of DPR rainfall rates. At KSHV radar site, for all cases, the MB is found to be positive indicating the GR rainfall estimates are low compared to

DPR products. This might be due to under performance of the dual-pol retrieval algorithm at this site. The overall poor performance of DPR rainfall products in convective case can be due several factors such as DPR attenuation error and non-uniform beam filling effect at Ku- and Ka-bands. Moreover, convective precipitation can be very localized resulting in precipitation cells smaller than the DPR horizontal resolution of 5 Km. This adds additional uncertainty in rainfall estimation.

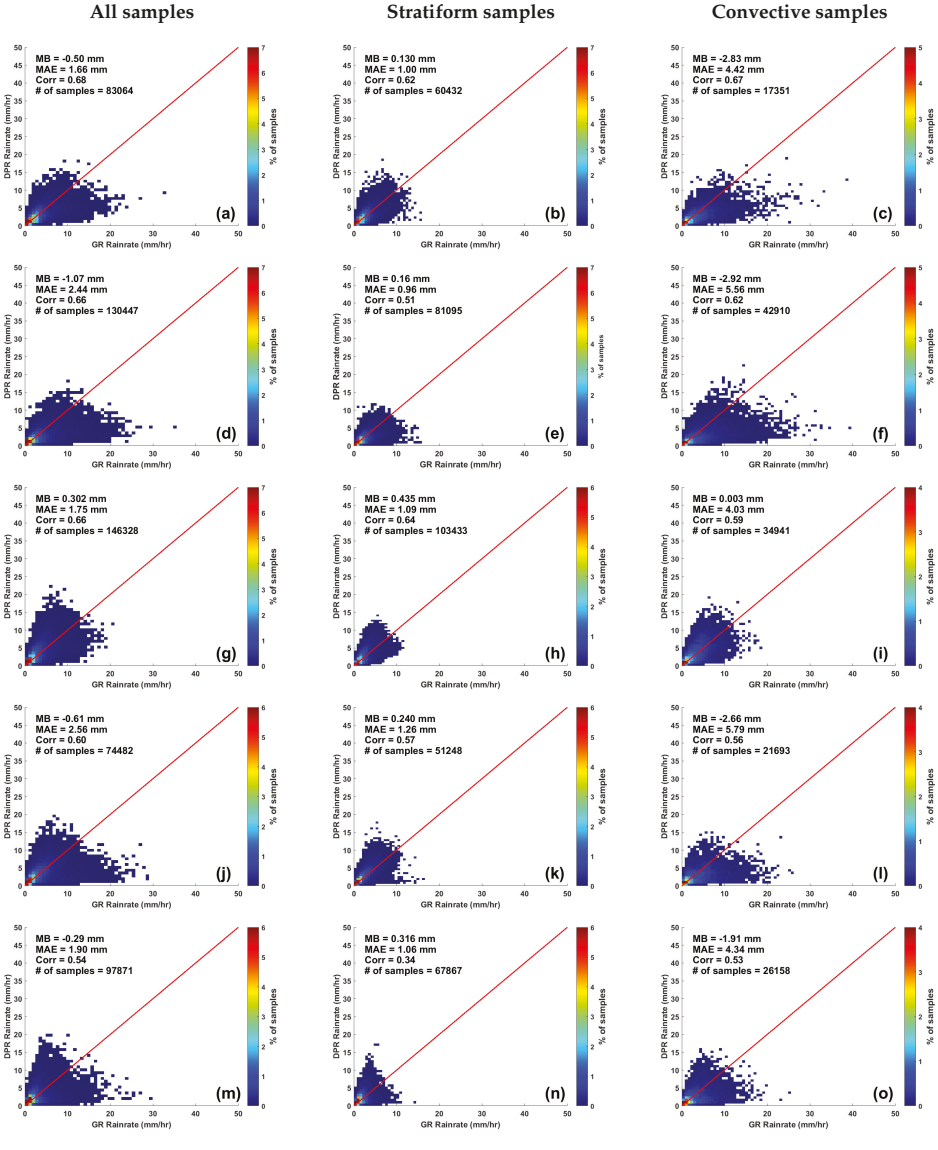


Figure 7. Scatter density plots of volume matched rainfall rate between GPM-DPR single frequency product and NEXRADs dual-pol product. Solid red line is 1:1 line. (**a**–**c**) KFWS radar; (**d**–**f**) KHGX radar; (**g**–**i**) KSHV radar; (**j**–**l**) KLIX radar; (**m**–**o**) KMLB radar.

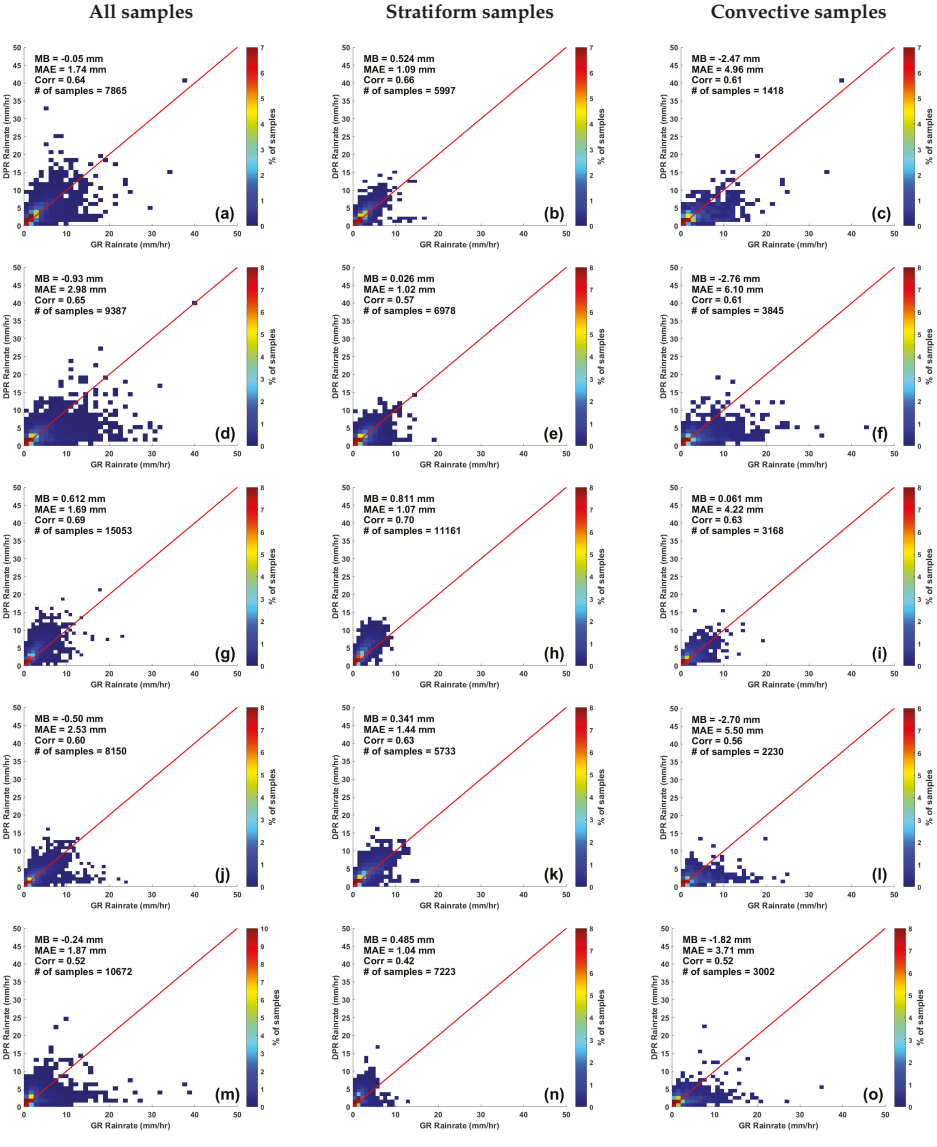


Figure 8. Same as Figure 7 but for GPM-DPR dual-frequency product.

Table 3. Table summarizing different error statistics for comparison of DPR single and dual-frequency rainfall rate product vs NEXRAD dual-pol product.

	All Samples						Stratiform Samples						Convective Samples					
	Single Frequency Product (mm/h)			Dual Frequency Product (mm/h)			Single Frequency Product (mm/h)			Dual Frequency Product (mm/h)			Single Frequency Product (mm/h)			Dual Frequency Product (mm/h)		
Radar	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**	**MB**	**MAE**	**Corr**
KFWS	−0.50	1.66	0.68	−0.05	1.74	0.64	0.13	1.00	0.62	0.52	1.09	0.66	−2.83	4.42	0.67	−2.47	4.96	0.61
KHGX	−1.07	2.44	0.66	−0.93	2.98	0.65	0.16	0.96	0.51	0.02	1.02	0.57	−2.92	5.56	0.62	−2.76	6.10	0.61
KSHV	0.30	1.75	0.66	0.61	1.69	0.69	0.43	1.09	0.64	0.81	1.07	0.7	0.00	4.03	0.59	0.06	4.22	0.63
KLIX	−0.61	2.56	0.60	−0.5	2.53	0.6	0.24	1.26	0.57	0.34	1.44	0.63	−2.66	5.79	0.56	−2.70	5.50	0.56
KMLB	−0.29	1.90	0.54	−0.24	1.87	0.52	0.31	1.06	0.34	0.48	1.04	0.42	−1.91	4.34	0.53	−1.82	3.71	0.52

4. Summary and Conclusions

In this study, GPM-DPR's reflectivity measurements at both Ku- and Ka-band and rainfall rate products are cross validated against 5 dual-pol ground-based radar from the GPM GV network. Ground radars located in the plains are chosen to avoid difficulties in precipitation observation over complex terrain. Over 250 GPM overpasses from April 2014 to June 2018 containing significant precipitation are studied. To reduce temporal mismatch, GR data is chosen such that the time difference between GR sweeps and the DPR overpass time is less than 5 min. Scattering simulation has also been performed to get an idea of ideal relationship between S-band, Ku-band, and Ka-band reflectivities in rain. The comparison procedure requires a volume matching technique to be applied on data from both GPM-DPR and ground radar. The methodology averages gates within a common volume at each DPR beam with GR sweeps. DPR data is only linearly averaged vertically along the beam whereas the GR data is inverse weighted averaged within DPR's footprint which is about 4.9 Km in diameter. Outcome of this methodology is matched DPR and GR samples located at each intersection points. The advantage of this technique over a common grid re-sampling are that there is no interpolation or extrapolation of the data. So, both DPR and GR have matched data taken at locations of actual observation.

Comparison study of matched DPR and GR reflectivity samples showed good agreement with correlation up to 0.9 at Ku-band and 0.85 at Ka-band. To better understand the correlation, samples are divided into stratiform rain and convective rain. The study revealed that stratiform samples are more correlated compared to convective samples. This is due to under performance of DPR attenuation correction algorithm in heavy precipitation. At all radar locations, mean biases for Ku-band is found to be positive while for Ka-band it is found to be negative. This is in well agreement to the simulated results. The MAEs are also within the difference limit found from simulations which is up to 3 dB for Ku-band and up to 5 dB for Ka-band. Though simulation results confirm the cause of difference in reflectivities, other factors such as attenuation correction, nonstandard GR beam refraction and errors in geometric match-up process should be considered as well. Other potential source of error is the time difference between DPR and GR samples. It can be reduced by applying advection correction to the GR data, but it has not been attempted in the present study. Next, absolute calibration of the 5 ground radars are tracked over the months using all GPM overpass data used in this study. Periods of relative stability in calibration have been observed with occasional variation in mean biases which can accounted for precipitation type variability across all seasons. Sudden jumps in MB is particularly noticeable in radars KSHV, KLIX and KMLB. The cause can be a combination of attenuation correction error and fewer GPM overpass for that specific month. Comparison of the volume matched rainfall rate samples revealed that the performance of DPR Ku-only product and dual-frequency product are similar for all the radars with correlation ranging from 0.52 to 0.69. The magnitude of error is higher for dual-frequency product compared to Ku-only product. It is also noticed that with higher rain rates the correlation of both Ku-only and dual-frequency product falls. Upon comparing samples classified by rain type, it is observed both DPR products greatly underestimate heavy convective rainfall by a margin of 5 to 6 mm/h at specific radar sites. In case of stratiform rain error magnitudes lying within 1.5 mm/h suggest far better performance of DPR products compared to convective rain. When considering all samples, the overall rainfall performance is found to be satisfactory. Although the GPM attenuation correction algorithm has performance issues and looking at reasonable agreement between reflectivities, it can be said that there are issues in the conversion of reflectivity into rainfall rate which could be improved to achieve more accurate rainfall estimation from GPM-DPR.

Author Contributions: Formal Analysis, Software, Writing and Original Draft Preparation was done by S.K.B. Conceptualization, Supervision, Validation, Editing and Review was done by V.C.

Funding: This project is funded by the NASA GPM program.

Acknowledgments: The authors acknowledge the data sources from the GPM program and the NOAA for the ground radar.

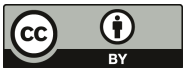
Conflicts of Interest: The authors declare no conflict of interest.

References

1. 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]
2. Liu, Z.; Ostrenga, D.; Teng, W.; Kempler, S. Tropical Rainfall Measuring Mission (TRMM) precipitation data and services for research and applications. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1317–1325. [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. Furukawa, K.; Nio, T.; Konishi, T.; Oki, R.; Masaki, T.; Kubota, T.; Iguchi, T.; Hanado, H. Current status of the dual-frequency precipitation radar on the global precipitation measurement core spacecraft. In *Sensors, Systems, and Next-Generation Satellites XIX*; International Society for Optics and Photonics: Bellingham, MA, USA, 2015; Volume 9639.
5. Senbokuva, Y.; Satoh, S.; Furukawa, K.; Koiima, M.; Hanado, H.; Takahashi, N.; Iguchi, T.; Nakamura, K. Development of the spaceborne dual-frequency precipitation radar for the Global Precipitation Measurement mission. In Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium, IGARSS'04, Anchorage, AK, USA, 20–24 September 2004; IEEE: Piscataway, NJ, USA, 2004; Volume 5.
6. Draper, D.W.; Newell, D.A.; Wentz, F.J.; Krimchansky, S.; Skofronick-Jackson, G.M. The global precipitation measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3452–3462. [CrossRef]
7. Liao, L.; Meneghini, R. A study on the feasibility of dual-wavelength radar for identification of hydrometeor phases. *J. Appl. Meteorol. Climatol.* **2011**, *50*, 449–456. [CrossRef]
8. Le, M.; Chandrasekar, V.; Biswas, S. An algorithm to identify surface snowfall from gpm dpr observations. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4059–4071. [CrossRef]
9. Chandrasekar, V.; Hou, A.; Smith, E.; Bringi, V.N.; Rutledge, S.A.; Gorgucci, E.; Petersen, W.A.; Jackson, G.S. Potential role of dual-polarization radar in the validation of satellite precipitation measurements: Rationale and opportunities. *Bull. Am. Meteorol. Soc.* **2008**, *89*, 1127–1146. [CrossRef]
10. Iguchi, T.; Seto, S.; Meneghini, R.; Yoshida, N.; Awaka, J.; Kubota, T. *GPM/DPR Level-2 Algorithm Theoretical Basis Document*; Tech. Rep.; NASA Goddard Space Flight Center: Greenbelt, MD, USA, 2010.
11. Iguchi, T.; Seto, S.; Meneghini, R.; Yoshida, N.; Awaka, J.; Kubota, T.; Kozu, T.; Chandra, V.; Le, M.; Liao, L.; et al. An overview of the precipitation retrieval algorithm for the dual-frequency precipitation radar (DPR) on the global precipitation measurement (GPM) mission's core satellite. In *Earth Observing Missions and Sensors: Development, Implementation, and Characterization II*; International Society for Optics and Photonics: Bellingham, MA, USA, 2012; Volume 8528.
12. Kubota, T.; Yoshida, N.; Urita, S.; Iguchi, T.; Seto, S.; Meneghini, R.; Awaka, J.; Hanado, H.; Kida, S.; Oki, R. Evaluation of precipitation estimates by at-launch codes of GPM/DPR algorithms using synthetic data from TRMM/PR observations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3931–3944. [CrossRef]
13. Toyoshima, K.; Masunaga, H.; Furuzawa, F.A. Early evaluation of Ku-and Ka-band sensitivities for the global precipitation measurement (GPM) dual-frequency precipitation radar (DPR). *Sola* **2015**, *11*, 14–17. [CrossRef]
14. Khajonrat, D.; Chandrasekar, V. Simulation of spaceborne radar observations of precipitation: Application to GPM-DPR. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2008, Boston, MA, USA, 7–11 July 2008; IEEE: Piscataway, NJ, USA, 2008; Volume 4.
15. Chandrasekar, V.; Khajonrat, D. Simulation of space-borne radar observations of precipitation at ku and ka band. Preprints. In Proceedings of the 34th Conference on Radar Meteorology, Williamsburg, VA, USA, 5–9 October 2009; Volume 7.
16. Speirs, P.; Gabella, M.; Berne, A. A comparison between the GPM dual-frequency precipitation radar and ground-based radar precipitation rate estimates in the Swiss Alps and Plateau. *J. Hydrometeorol.* **2017**, *18*, 1247–1269. [CrossRef]
17. Gabella, M.; Speirs, P.; Hamann, U.; Germann, U.; Berne, A. Measurement of Precipitation in the Alps Using Dual-Polarization C-Band Ground-Based Radars, the GPM Spaceborne Ku-Band Radar, and Rain Gauges. *Remote Sens.* **2017**, *9*, 1147. [CrossRef]

18. Biswas, S.K.; Chandrasekar, V. Cross validation of observations from GPM dual-frequency precipitation radar with S-band ground radar measurements over the Dallas—Fort worth region. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017.
19. Warren, R.A.; Protat, A.; Siems, S.T.; Ramsay, H.A.; Louf, V.; Manton, M.J.; Kane, T.A. Calibrating ground-based radars against TRMM and GPM. *J. Atmos. Ocean. Technol.* **2018**, *35*, 323–346. [CrossRef]
20. Crisologo, I.; Warren, R.A.; Mühlbauer, K.; Heistermann, M. Enhancing the consistency of spaceborne and ground-based radar comparisons by using beam blockage fraction as a quality filter. *Atmos. Meas. Tech.* **2018**, *11*, 5223–5236. [CrossRef]
21. Panegrossi, G.; Casella, D.; Dietrich, S.; Marra, A.C.; Sanò, 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]
22. 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]
23. Derin, Y.; Anagnostou, E.; Anagnostou, M.N.; Kalogiros, J.; Casella, D.; Marra, A.C.; Panegrossi, G.; Sanò, 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]
24. Anagnostou, M.N.; Kalogiros, J.; Nikolopoulos, E.; Derin, Y.; Anagnostou, E.N.; Borga, M. Satellite rainfall error analysis with the use of high-resolution X-band dual-polarization radar observations over the Italian Alps. In *Perspectives on Atmospheric Sciences*; Springer: Cham, Switzerland, 2017; pp. 279–286.
25. Schwaller, M.R.; Morris, K.R. A ground validation network for the global precipitation measurement mission. *J. Atmos. Ocean. Technol.* **2011**, *28*, 301–319. [CrossRef]
26. NASA/JAXA. Precipitation Measurement Missions: Global Precipitation Mission (GPM) Ground Validation System. 2015. Available online: https://pmm.nasa.gov/sites/default/files/document_files/Val_Network_Users_Guide_Vol_2_Nov2015.pdf (accessed on 31 July 2018).
27. Seto, S.; Iguchi, T. Intercomparison of attenuation correction methods for the GPM dual-frequency precipitation radar. *J. Atmos. Ocean. Technol.* **2015**, *32*, 915–926. [CrossRef]
28. Bolen, S.M.; Chandrasekar, V. Methodology for aligning and comparing spaceborne radar and ground-based radar observations. *J. Atmos. Ocean. Technol.* **2003**, *20*, 647–659. [CrossRef]
29. Le, M.; Chandrasekar, V. Raindrop size distribution retrieval from dual-frequency and dual-polarization radar. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1748–1758. [CrossRef]
30. Le, M.; Chandrasekar, V.; Biswas, S. Evaluation and Validation of GPM Dual-Frequency Classification Module after Launch. *J. Atmos. Ocean. Technol.* **2016**, *33*, 2699–2716. [CrossRef]
31. Biswas, S.K.; Le, M.; Chandrasekar, V. Identification of Snow from GPM-DPR observations and cross validation with S-Band Ground Radar dual polarization measurements. In Proceedings of the 2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS), Montreal, QC, Canada, 19–26 August 2017.
32. NASA/JAXA. Global Precipitation Measurement Precipitation Processing System: File Specification for GPM Products. 2018. Available online: https://pps.gsfc.nasa.gov/Documents/filespec.GPM_V5.pdf (accessed on 31 July 2018).
33. NASA/JAXA. Global Precipitation Measurement Precipitation Processing System: Release Notes for the DPR V5 Level 2 Products. 2017. Available online: https://pps.gsfc.nasa.gov/Documents/V05ReleaseNotes/Caveats_DPRL2_productV05.pdf (accessed on 31 July 2018).
34. Klazura, G.E.; Imy, D.A. A description of the initial set of analysis products available from the NEXRAD WSR-88D system. *Bull. Am. Meteorol. Soc.* **1993**, *74*, 1293–1312. [CrossRef]
35. 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]
36. Ryzhkov, A.V.; Giangrande, S.E.; Melnikov, V.M.; Schuur, T.J. Calibration issues of dual-polarization radar measurements. *J. Atmos. Ocean. Technol.* **2005**, *22*, 1138–1155. [CrossRef]
37. Waterman, P.C. Symmetry, unitarity, and geometry in electromagnetic scattering. *Phys. Rev. D* **1971**, *3*, 825. [CrossRef]

38. 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]
39. 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]
40. Baldini, L.; Chandrasekar, V.; Moisseev, D. Microwave radar signatures of precipitation from S band to Ka band: application to GPM mission. *Eur. J. Remote Sens.* **2012**, *45*, 75–88. [CrossRef]
41. Bolen, S.M.; Chandrasekar, V. Quantitative Cross Validation of Space-Based and Ground-Based Radar Observations. *J. Appl. Meteorol.* **2000**, *39*, 2071–2079.<2071:QCVOSB>2.0.CO;2. [CrossRef]
42. Anagnostou, E.N.; Morales, C.A.; Dinku, T. The use of TRMM precipitation radar observations in determining ground radar calibration biases. *J. Atmos. Ocean. Technol.* **2001**, *18*, 616–628. [CrossRef]
43. Liao, L.; Meneghini, R.; Iguchi, T. Comparisons of rain rate and reflectivity factor derived from the TRMM precipitation radar and the WSR-88D over the Melbourne, Florida, site. *J. Atmos. Ocean. Technol.* **2001**, *18*, 1959–1974. [CrossRef]
44. Barnes, S.L. A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteorol.* **1964**, *3*, 396–409. [CrossRef]
45. Meneghini, R.; Iguchi, T.; Kozu, T.; Liao, L.; Okamoto, K.I.; Jones, J.A.; Kwiatkowski, J. Use of the surface reference technique for path attenuation estimates from the TRMM precipitation radar. *J. Appl. Meteorol.* **2000**, *39*, 2053–2070. [CrossRef]
46. Le, M.; Chandrasekar, V. Precipitation type classification method for dual-frequency precipitation radar (DPR) onboard the GPM. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1784–1790. [CrossRef]
47. Iguchi, T.; Kozu, T.; Meneghini, R.; Awaka, J.; Okamoto, K.I. Rain-profiling algorithm for the TRMM precipitation radar. *J. Appl. Meteorol.* **2000**, *39*, 2038–2052. [CrossRef]
48. Konrad, C.E.; Perry, L.B. Relationships between tropical cyclones and heavy rainfall in the Carolina region of the USA. *Int. J. Climatol.* **2010**, *30*, 522–534. [CrossRef]
49. Chen, H.; Chandrasekar, V. The quantitative precipitation estimation system for Dallas–Fort Worth (DFW) urban remote sensing network. *J. Hydrol.* **2015**, *531*, 259–271. [CrossRef]
50. Omranian, E.; Sharif, H.; Tavakoly, A. How well can global precipitation measurement (GPM) capture hurricanes? Case study: Hurricane Harvey. *Remote Sens.* **2018**, *10*, 1150. [CrossRef]
51. Omranian, E.; Sharif, H.O. Evaluation of the Global Precipitation Measurement (GPM) Satellite Rainfall Products over the Lower Colorado River Basin, Texas. *JAWRA J. Am. Water Resour. Assoc.* **2018**, *54*, 882–898. [CrossRef]
52. 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]



remote sensing

MDPI

Article

Assessment of Satellite and Radar Quantitative Precipitation Estimates for Real Time Monitoring of Meteorological Extremes Over the Southeast of the Iberian Peninsula

Fulgencio Cánovas-García [1,*,†], Sandra García-Galiano [1,†] and Francisco Alonso-Sarría [2,†]

[1] Unidad Predepartamental de Ingeniería Civil, Universidad Politécnica de Cartagena, Paseo Alfonso XIII, 52, 30203 Cartagena, Spain; sandra.garcia@upct.es

[2] Instituto Universitario del Agua y del Medio Ambiente, Universidad de Murcia, Edificio D, Campus de Espinardo, s/n, 30100 Murcia, Spain; alonsarp@um.es

* Correspondence: fulgencio.canovas@upct.es

† These authors contributed equally to this work.

Received: 9 May 2018; Accepted: 20 June 2018; Published: 27 June 2018

Abstract: Quantitative Precipitation Estimates (QPEs) obtained from remote sensing or ground-based radars could complement or even be an alternative to rain gauge readings. However, to be used in operational applications, a validation process has to be carried out, usually by comparing their estimates with those of a rain gauges network. In this paper, the accuracy of three QPEs are evaluated for three extreme precipitation events in the last decade in the southeast of the Iberian Peninsula. The first QPE is PERSIANN-CCS (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Cloud Classification System) , a satellite-based QPE. The second and the third are QPEs from a meteorological radar with Doppler capabilities that works in the C band. Pixel-to-point comparisons are made between the values offered by the QPEs and those obtained by two networks of rain gauges. The results obtained indicate that all the QPEs were well below the rain gauge values in extreme rainfall time slots. There seems to be a weak linear association between the value of the discrepancies and the precipitation value of the QPEs. The main conclusion, assuming the information from the rain gauges as ground truth, is that neither PERSIANN-CCS nor radar, without empirical calibration, are acceptable QPEs for the real-time monitoring of meteorological extremes in the southeast of the Iberian Peninsula.

Keywords: quantitative precipitation estimates; validation; PERSIANN-CCS; meteorological radar; satellite rainfall estimates

1. Introduction

Precipitation is a highly relevant feature in Earth sciences. Precipitation estimations with good spatial and temporal resolution are important in hydrology, climate research [1], ecology and meteorology. Precipitation estimations might be the most important meteorological input for calibrating and using hydrological and ecological models [2].

Meteorological radars have been used for flood warnings, but they are very expensive to set up and maintain, and their coverage is limited in mountainous areas [3]. In fact, the first operational network of meteorological radars in tropical mountain areas has only been operational for a few years [4]. These authors and Nikolopoulos et al., 2013 [5] also identified the difficulty of maintaining radar networks in developing countries. This lack of adequate data for precipitation monitoring limits the scope for hydro-meteorological research and the use of physical or statistical models for water resources management [6].

Rain gauge networks have been used as primary source of rainfall measurements for over a century. However, while these devices provide direct and accurate (relative to other sensors) rainfall measurements, they are associated with small sampling areas [5]. Another problem arises when using rain gauges as a source of information of precipitation: rain gauges perform point specific measurements and, even in very dense networks, it may not be possible to capture the spatial variability of precipitation, especially when working at subhour scales [7] or when dealing with very localised convective or orographic precipitation.

The Quantitative Precipitation Estimates (QPEs) obtained by remote sensing can complement or be considered an alternative to rain gauge measurements. Satellite-based QPE are valuable continuous records on several temporal [8] and spatial scales and can provide acceptable good estimations of precipitation in "un-gauged" regions, such as oceans, hard-to-reach mountainous areas or deserts [2]. We agree with Zambrano-Bigiarini et al. (2017) [9] that satellite-based QPEs provide an unprecedented opportunity for a wide variety of meteorological and hydrological applications.

In recent decades, the frequency of flood disasters in European Union (EU) has increased [10]. In Spain, floods are the natural hazard with the greatest territorial impact and are responsible for great socio-economic losses [11]. Spain is also the EU country most affected by flash floods [10]. In the southeast of the Iberian Peninsula these phenomena have caused a very high number of deaths and millions of economic losses. Recent examples are the three events analysed in Section 2.2.

It is not possible for all small and medium sized basins that may produce a flooding event to measure runoff in the channels, so precipitation intensity remains the most used source of information during extreme weather events. Therefore, regardless of the methodology or type of modelling used, warning accuracy will depend on the accuracy of the precipitation estimation [5].

Serrano Notivoli et al. (2017) [12] developed a high resolution daily precipitation grid for the whole of Spain from rain gauge data. They also analysed the estimation uncertainties, concluding that the highest uncertainty values appear in SE Spain, the study area of this paper. The reason is that most of the precipitation in this area is produced by convective systems generated over the Mediterranean Sea interacting with Potential Vorticity Streamers (PVS) or cut-offs in the higher troposphere [13]. In such conditions, the entrance of wet air masses to the land is driven by the relief pattern and the wind direction. The interaction of both factors produces highly localised upwinds and a very irregular precipitation pattern that are very difficult to estimate by interpolation with daily rain gauge data alone. This is obviously a considerable handicap for hydrological forecasting. This concentration and variability justifies any attempt to obtain better rainfall estimations using remote sensing products.

Satellite-based QPEs have not been well integrated into operational and decision-making applications because of the lack of rigorous validation and uncertainty analysis [14]. Hong et al. (2007) [15] pointed out that the strong spatio-temporal variability of precipitation makes it necessary to rigorously assess QPE accuracy before the estimates can be used with confidence. In a recent paper, Zambrano-Bigiarini et al. (2017) [9] claimed that no satellite-based QPE can be generally considered more accurate than any others on a daily scale, and that accuracy must be assessed on a case-by-case basis in the study area in question.

Accuracy estimations of global products such as PERSIANN-CCS (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System) in areas where rain gauge data is available might be useful to estimate if they are accurate enough to be used in areas where rigurous validation is impossible due to the absence of rain gauge information with which to compare satellite-based QPEs.

The aim of this research was to evaluate the accuracy of two kinds of QPEs during heavy rainfall events. The intention was also to ascertain whether any of the QPEs could be regarded as a good substitute for the information provided by a rain gauge network if it malfunctions.

We have compared precipitation measurements of two rain gauge networks with those estimated by a satellite-based QPE (PERSIANN-CCS) and two ground-based QPEs (meteorological radar) during the three most important heavy rainfall episodes recorded during the last ten years in the southeast

of the Iberian Peninsula (Spain). Comparisons are made with aggregated data, using statistics for the entire study area and for the most affected river basins, on the one hand, and in a disaggregated manner, doing the same for the rain gauge that recorded the greatest accumulations and intensities of precipitation, on the other. The comparisons are made on a spatial basis, using interpolation to map the differences in accumulated precipitation between the three QPEs and the rain gauges on the one hand, and between the QPEs, on the other. Comparisons are also made with a temporal perspective using statistics on the agreement among precipitation sources calculated on an hourly scale.

2. Materials and Methods

2.1. Study Area

The research was carried out in the terrestrial portion of the Segura River Basin District, located in south-eastern Spain (Figure 1), with a surface area of 18,740 km^2 covering the Segura River basin and other small coastal basins [16]. It is a territory with scarce and irregular rainfall, high temperatures and a high annual number of hours of sun. There is a NW-SE precipitation gradient that ranges from approximately 1000 mm/year in the headwaters of the Segura river to less than 300 mm/year in the coastal zone [17]. With the exception of the Segura river, the most common channels are ramblas (ephemeral rivers), frequently responsible for flash floods [11]. Although it is a small river basin district, its management is quite complex as a result of having water resources from different sources (surface and groundwater resources, desalination, transfers and reuse) and multiple uses that compete for the scarce water resources [18].

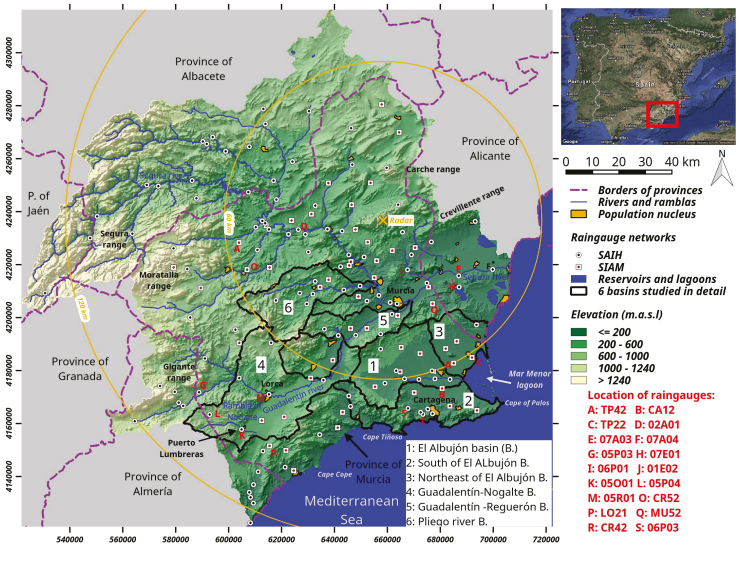


Figure 1. Map of the study area, the Segura River Basin District. Coordinate Reference System is ETRS89 with projected coordinates (EPSG—European Petroleum Survey Group—code: 25830).

2.2. Rainfall Events

Three very intense precipitation events, in the last decade were identified and characterized. All off them were convective in nature. The first one took place on 27–28 September 2009 and affected the Campo of Cartagena area. The rain gauge 06P03, located about 5 km NW of La Vaguada, recorded 199 mm in 29 h and the CA12 rain gauge, located in La Palma, recorded 268 mm in 30 h. The two

mentioned rain gauges were located in the municipality of Cartagena, and the rainfall produced several flash floods in the Campo de Cartagena basins (basins number 1, 2 and 3 in Figure 1).

The second event occurred during the first half of 28 September, 2012. The rain gauge 05P03, located in the foothills of Sierra del Gigante (Municipality of Lorca), recorded 124 mm in five hours and the 05O01 rain gauge, located in Puerto Lumbreras, recorded 153 mm in six hours. Massive flash floods were registered in the Nogalte (Figure 2) and Guadalentín basins (basin number 4 in Figure 1).

Figure 2. The rambla of Nogalte is normally dry (**a**) and only runs water when there is heavy rainfall in the upper basin as on the 28 September 2012 event (**b**). Sources: (**a**) [19]; (**b**) CHS (Segura Basin Hydrological Confederation) , 2013.

The most recent event was registered between 17 and 19 December 2016, and was the most important meteorological episode of the whole year for the Iberian Peninsula. The highest rainfall intensities were registered in Campo of Cartagena and the east coast of the study area, where the rainfall exceeded 50 mm in one hour in the Torre Pacheco (TP42) and San Javier (TP22) rain gauges. Areas close to the Mar Menor lagoon, which environment has undergone a strong process of global transformation due to tourist [20] and farming activity, were severely affected by the storm and the consequent flooding.

The three analyzed events present a high return period. The event of September 2009 corresponds to 200-years return period. Moreover, both events of September 2012 and December 2016 correspond to 500-years return period. The estimations were based on Ministerio de Fomento (1999) [21] by applying the SQRT-ETmax cumulative distribution function.

2.3. Satellite-Based Quantitative Precipitation Estimates : PERSIANN-CCS

PERSIANN-CCS [1] is a system based on satellite imagery and pattern recognition techniques applied for the automatic classification of several types of clouds in order to estimate the rainfall in each pixel.

PERSIANN-CCS produces QPEs with a time resolution of one hour, a spatial resolution of 0.04° ($\approx 4 \times 4$ km^2 in Spanish latitudes), with near to global coverage (between 60°N and 60°S) and a lag time of approximately one hour (near-real time). PERSIANN-CCS is a product purely derived from satellite observations, i.e., it is not calibrated using ground observations [22]. Since PERSIANN-CCS is available in near-real time, it is suitable for use in flood warning and management applications [3,23], especially in large river systems such as the Segura or Guadalentín river basins where the one-hour temporal resolution has little impact on hydrological analysis compared to smaller flash floods prone basins [24]. Very promising advances have been made in the calibration of PERSIANN-CCS with data from other satellite data [25], but, unfortunately, these products are not available in near-real time.

2.4. Ground-Based Quantitative Precipitation Estimates : Meteorological Radar

Radar data from the Spanish Meteorology Service (AEMET) has been used. This apparatus operates in the C band (5.6 GHz) and is equipped with Doppler capability. It currently provides data

for a circle of 240 km radius (long range mode) and provides images in Cartesian local projection according to Lambert's conformal conical centered on the radar (38.27°N, 1.19°W). Each image consists of 480 × 480 pixels and a spatial resolution of 1 × 1 km^2 when operating in long-range mode. When working in short-range mode, it provides information for a circle area of 120 km radius and a spatial resolution of 500 × 500 m^2. Hourly accumulated precipitation data, the QPE with highest temporal resolution offered by AEMET, were used in this work. Meteorological radar data is not freely accessible. An official request to AEMET is required and the price is 0.51 euros/image plus taxes, except when a discount for scientific research is approved by AEMET.

Although operational QPEs provided by AEMET are used in this research, we will briefly describe how the radar reflectivity is transformed to rainfall intensity. The reflectivity images have a temporal resolution of 10 min and store the reflectivity Z in dB (ZdB), this is transformed to reflectivity (Z) in $mm^6\ mm^{-3}$ using the equation $ZdB = 10\ \log_{10}\ Z$, after that the rainfall intensity in mm h^{-1} is obtained applying Marshall-Palmer's Z-R ratio [26] (p. 178) $Z = 200\ R^{1.6}$. The QPE CAPPI (constant altitude plan position indicator) with a temporal resolution of 1 hour is obtained after averaging the 6 corresponding intensities.

For the 2009 episode, hourly accumulation data, based on CAPPI , are available for working in short range mode. No meteorological radar data are available for the 2012 episode due to a power failure because of the storm . Finally, for the 2016 episode, an hourly QPE called SRI (surface rainfall intensity) was used. This is an improved product that takes into account the nature of the precipitation (convective or stratiform) before applying, or not, a correction for the vertical reflectivity profile. The main advantage of radar QPEs over a satellite-based QPEs is a lag time of about 7 min from the end of the accumulation period, as shown in the image metadata.

2.5. Rain Gauges

PERSIANN-CCS and radar data were compared with two rain gauge networks. The SIAM (Agroclimatic Information Service of Murcia) network consists of several automatic rain gauges (45 in the 2009 event, 44 in the 2012 event and 47 in the 2016 event) with a temporal resolution of one hour. The second network is the SAIH-Segura (Automatic Hydrological Information System of the Segura River Basin) operated by the Water Authority (Segura Basin Hydrological Confederation, CHS). It had 64 rain gauges operative for the 2009 event, 66 for the 2012 event, and 106 for the 2016 event. The time resolution of this network is 5 min. The CRS of the two rain gauge networks is ETRS89/UTM zone 30N (EPSG code: 25830). Both networks were joined to obtain a more dense network. In 2009 and 2012 there was a rain gauge for every 173 km^2, for the 2016 event the density increased to a rain gauge every 124.5 km^2.

The data from SIAM network are evaluated and validated internally before being made available to the public. SAIH-Segura network data for the 2009 and 2012 events are reported in the system as "filtered and consolidated". Only for the 2016 event did the data appear at the time of the download as "provisional, obtained in real time without checking". Regardless of this, the precipitation data have been analyzed to eliminate possible erroneous values prior to comparison with the QPEs using the methodology proposed by Velasco-Forero et al. (2009) [27], which is: Rain gauges with cumulative precipitation lower than 1.5 mm from the entire event were discarded from the analysis when the cumulative radar precipitation was more than 10 mm. In addition, long periods of inactivity of the rain gauges have been monitored to ensure that they correspond to the same pattern in the radar data. After the corresponding analyses, one rain gauge was removed from the SIAM network in the 2009 event and another one from the SAIH-Segura network in the 2016 event.

2.6. Assesssment of Quantitative Precipitation Estimates

As proposed by other authors [9,28–30] a point-to-pixel analysis was used to compare rain gauge data with the QPEs. The rain gauge layers were reprojected to the CRS of the QPE to avoid uncertainties associated with the resampling of the pixels, and when more than one rain gauge intersected with

a QPE pixel, the mean value of the rain gauges was used for the comparison. When this is the case, the number of pixels available for comparison (second column of the Table 1) is not the same as the number of rain gauges.

To implement a point-to-pixel comparison is a difficult task because of several uncertainties to be taken into account when evaluating the statistics derived from the comparison. For example, the spatial support of QPE and rain gauges is different. The rain gauge entrance (in all cases less than 0.05 m^2) can be approximated to a point measurement, while the values stored in the pixels of the QPEs correspond to averages over the volume of a grid cell [7]. This causes a smoothing of the QPEs' values compared to the punctual measurements of the rain gauges [31]. Therefore, the differences in the values of rain gauges and QPE pixels are not only due to errors in the QPEs, but also to differences in spatial support [7]. This problem increases if the spatial resolution of the QPE is smaller, so the problem is much more serious for PERSIANN-CCS than for both radar images.

In addition, the finescale variability of precipitation even at short distances, especially with convective precipitation, which cannot be represented by a dispersed network of rain gauges [15], introduces uncertainty into the precipitation values averaged over large areas.

Such problems require careful interpretation of the observed differences when a point-to-pixel comparison is applied. According to Schiemann et al. (2011) [7], accepting the assumption that the above effects lead to a random component in the pluviometer-QPE differences, comparisons made for a large number of rain gauges (such as that used in this research) can provide some guidance concerning the accuracy obtained by different QPEs. For these reasons, the word "error" is avoided as the word "difference" is considered more appropriate; although terms such as overestimate or underestimate (always using the rain gauge value as the baseline) are introduced with the intention of simplifying the text as much as possible.

The statistics calculated to report the additive component of the differences between rain gauges and QPEs were: $diff_i = qpe_i - gau_i$, where gau_i was the measured precipitation for a given rain gauge during a given time interval and qpe_i was the estimated precipitation for the QPE at the pixel intersecting with the location of the previous rain gauge for the same time interval; the average rainfall of rain gauges as $\overline{gau} = \frac{1}{n}\sum gau_i$; the average precipitation of the QPE, $\overline{qpe} = \frac{1}{n}\sum qpe_i$; the root mean square difference as $RMSD = \sqrt{\frac{1}{n}\sum diff_i^2}$; the mean absolute difference as $MAD = |\frac{1}{n}|diff_i|$; the bias as $\frac{1}{n}\sum diff_i$; Pearson's linear correlation coefficient [32] (p. 134); the relative MAD as $rMAD = \frac{MAD}{\overline{gau}}$ and the relative RMSD as $rRMSD = \frac{RMSD}{\overline{gau}}$.

In very intense but very localised precipitations, the use of these statistics might produce misleading results since the values close to zero have a downward influence on the value of the statistic. In research such as this, the use of conditional statistics is desirable. In our case the condition was that either the rain measure (gau_i) or the QPE estimate (qpe_i) was equal to or greater than 1 mm/h. Conditional statistics are represented in this work with an asterisk in front of the statistic's name, so the conditional MAD is *MAD.

In addition to the aforementioned statistics, others, related to the multiplicative component of the differences, were calculated. This is the main component in many of the errors present in meteorological radar QPEs [33]. These statistics are usually calculated only with the cases in which both sources register precipitation. A 0.2 mm/h threshold, the lowest resolution of all the sensors used (some SIAM network rain gauges), was used. For each episode the bias in dB was calculated as $bias = 10\log_{10}\left(\frac{\sum qpe_i}{\sum gau_i}\right)$ [34]. This statistic expresses the overall agreement between QPEs estimates and ground truth given that both instruments register precipitation [35]. The scatter was also calculated as $10\log_{10}\left(\frac{|Q_{16}Diff - Q_{84}Diff|}{2}\right)$ where $Q_{16}Diff$ refers to the 16% percentiles of the $Diff$ distribution. The scatter refers to the spread of hourly QPE-gauge ratios when pooling all rain hours and gauges together [36]. Considering that this statistic is not very sensitive to outliers, it should be interpreted with caution when studying precipitation associated with high return period events.

Different comparisons were carried out to determine the degree of agreement between the three sources of information. A comparison of the precipitation values at a given location was made (Figures 9 and 10). A comparison of statistics has also been made for the course of the three storms in different spatial areas (Figures 3–5). This allowed us to identify the QPE whose rainfall values were closest to those of the rain gauges and to relate the observed differences with the precipitation intensities measured by the rain gauges. An analysis of the spatial distribution of the total accumulation of events has been carried out to compare the QPEs and identify, for each of them, the places where the greatest differences with respect to rain gauges occurred. Interpolation techniques were used for this purpose. Finally, a pixel-by-pixel comparison, as proposed by Nguyen et al. (2015) [24], between the radar and PERSIANN-CCS was carried out by generating a mapping of statistics and linear association between the QPEs for the 2009 and 2016 events (Figure 8).

3. Results

3.1. Statistics

Table 1 shows the statistics of the degree of agreement and linear association among rain gauges and QPEs. The number of pixels for comparison may not be equal for the same event in the two QPEs analysed due to their different spatial resolutions. The observations column indicates the number of pairs of precipitation values taken into account to calculate the statistics. In all the events, the number of non-conditional observations in the study area more than doubles the conditional observations. These large differences in extreme precipitation events reflect their highly localised character. Focusing on the basin with the greatest accumulation of precipitation in the 2009 event (647 km^2), the number of conditional observations is approximately 72 percent less than the number of non-conditional observations. In the 2012 event, the percentage of conditional observations is less than 50 in the sub-basin with the greatest accumulation of precipitation.

In the September 2009 event (Table 1a), the RMSD of the entire study area is moderate, 3.5 mm/h for PERSIANN-CCS and 2.9 mm/h for radar. However, this value is partly due to the high number of hours in which both PERSIANN-CCS and radar have precipitation values below 1. The conditional RMSD values increase to 5.2 and 4.9 mm/h, respectively. The relative values of RMSD indicate that the statistic is 1.8 times the mean precipitation value of PERSIANN-CCS and 1.4 times that of radar. Both PERSIANN-CCS and radar show worse agreements with the rain gauges located in the basins most affected by the storm. For the whole study area, the *bias values of PERSIANN-CCS indicate that the underestimated values compensate for the overestimated values. Both bias and *bias values are larger with radar for the whole study area and sub-basins. Correlation values, on the other hand, both conditional and non-conditional, are larger with radar in all cases.

No radar data were available for the 2012 event (Table 1b). PERSIANN-CCS statistics shows some degree of agreement with those of the 2009 event. The *RMSD is larger and this increase is proportional to the variation in precipitation intensity, which in the 2012 event was greater. *rRMSD and *rMAD values remain constant. In the 2012 event bias and *bias values indicate that PERSIANN-CCS as a whole underestimates precipitation and the correlation values indicate that the linear association is not strong.

For the 2016 event (Table 1c) another QPE from the same radar (SRI) was used. In a first analysis the values of *RMSD are similar to those of the 2009 episode. Significant changes in statistics can be seen when moving from conditional to non-conditional statistics. In the two QPEs, precipitation tends to be underestimated, the values for the whole basin being similar in both; however, if the most affected basins are taken into account, in two of them (Rambla of El Albujón basin -Al.- and Basin located at Northeast of rambla of El Albujón -NE. Al-) the *bias is unfavourable to radar and in the third the values are the same (Subasin of Guadalentín river where the Reguerón channel is located -G. Re.-). With respect to the correlation of the whole area, radar shows values that are closer to those of the rain gauges, but if the three basins analyzed are considered, the situation is variable. For this

episode and according to the data provided by this table and from the point of view of the additive component of the differences, it cannot be said that either QPE gives results that are more similar to the rain gauge results measurements.

With respect to the multiplicative components of the differences (Table 2) bias of PERSIANN-CCS is close to 0 especially in the 2009 event, whereas in the two events in which radar is available is much lower than in this one. The values of radar bias are quite high. The scatter values inform us that in the QPEs, and for all events, there is a very high dispersion, in both cases higher in PERSIANN-CCS and with the highest value in the 2012 event.

Table 1. Statistics of agreement and linear association between rain gauges and the corresponding pixels of PERSIANN-CCS and radar (additive component of the differences), for the three events in the basins most affected by each event.

(a) 27th and 28th September 2009 Event														
PERSIANN-CCS, Spatial Resolution ≈4 × 4 km²														
Name	**No. pix**	**Obs.**	***Obs.**	$\overline{qpe}$	$*\overline{qpe}$	$\overline{gau}$	$*\overline{gau}$	**RMSD**	***RMSD**	**rRMSD**	***rRMSD**	**Bias**	***Bias**	**R**	***R**
S. Al	5	220	158	2.9	4.0	5.0	6.9	8.2	9.6	1.6	1.4	−2.1	−2.9	0.34	0.18
Al	6	264	193	3.0	4.0	2.9	3.9	5.1	6.0	1.8	1.5	0.1	0.1	0.22	0.04
NE. Al	5	220	155	3.2	4.5	2.6	3.7	5.6	6.7	2.1	1.8	0.6	0.8	0.07	−0.17
Total	106	4664	2060	1.4	3.0	1.4	2.9	3.5	5.2	2.6	1.8	0.0	0.1	0.31	0.08
Radar, CAPPI Product in Short Range Mode, Spatial Resolution ≈500 × 500 m²														
Name	**No. pix**	**Obs.**	***Obs.**	$\overline{qpe}$	$*\overline{qpe}$	$\overline{gau}$	$*\overline{gau}$	**RMSD**	***RMSD**	**rRMSD**	***rRMSD**	**Bias**	***Bias**	**R**	***R**
S. Al	5	220	159	2.5	3.4	5.0	6.9	7.9	9.3	1.6	1.4	−2.5	−3.4	0.46	0.34
Al	6	264	168	1.8	2.8	2.9	4.5	4.2	5.2	1.4	1.2	−1.1	−1.7	0.49	0.31
NE. Al	5	220	110	1.2	2.3	2.6	5.1	4.6	6.5	1.8	1.3	−1.4	−2.9	0.29	−0.06
Total	108	4752	1664	0.8	2.1	1.3	3.6	2.9	4.9	2.2	1.4	−0.5	−1.5	0.50	0.28
(b) 27th and 28th Septemner 2012 Event														
PERSIANN-CCS, Spatial Resolution ≈4 × 4 km²														
Name	**No. pix**	**Obs.**	***Obs.**	$\overline{qpe}$	$*\overline{qpe}$	$\overline{gau}$	$*\overline{gau}$	**RMSD**	***RMSD**	**rRMSD**	***rRMSD**	**Bias**	***Bias**	**R**	***R**
G-No	8	216	97	3.0	6.6	4.2	9.4	11.0	16.4	2.6	1.8	−1.2	−2.8	0.32	0.13
R. Pli	6	162	92	2.5	4.4	3.8	6.6	10.0	13.2	2.6	2.0	−1.3	−2.1	0.08	−0.11
Total	107	2889	1379	2.0	4.1	2.7	5.6	7.7	11.1	2.8	2.0	−0.7	−1.5	0.18	−0.02
(c) 17–19 December 2016 Event														
PERSIANN-CCS, Spatial Resolution ≈4 × 4 km²														
Name	**No. pix**	**Obs.**	***Obs.**	$\overline{qpe}$	$*\overline{qpe}$	$\overline{gau}$	$*\overline{gau}$	**RMSD**	***RMSD**	**rRMSD**	***rRMSD**	**Bias**	***Bias**	**R**	***R**
Al	9	603	330	1.5	2.8	2.5	4.4	4.4	5.9	1.7	1.3	−1.0	−1.7	0.32	0.12
NE. Al	5	335	181	1.6	3.0	3.1	5.7	5.2	7.1	1.7	1.3	−1.5	−2.7	0.51	0.33
G-Re	7	469	282	1.4	2.3	2.6	4.2	3.5	4.5	1.3	1.1	−1.3	−2.0	0.41	0.19
Total	140	9380	4425	1.0	2.1	1.7	3.3	3.1	4.5	1.9	1.4	−0.6	−1.2	0.34	0.13
Radar, SRI Product in Long Range Mode, Spatial Resolution ≈1 × 1 km²														
Name	**No. pix**	**Obs.**	***Obs.**	$\overline{qpe}$	$*\overline{qpe}$	$\overline{gau}$	$*\overline{gau}$	**RMSD**	***RMSD**	**rRMSD**	***rRMSD**	**Bias**	***Bias**	**R**	***R**
Al	11	737	417	0.8	1.2	2.4	4.1	4.2	5.6	1.8	1.4	−1.6	−2.9	0.29	0.07
NE. Al	5	335	178	1.0	1.7	3.1	5.8	5.8	8.0	1.9	1.4	−2.1	−4.1	0.42	0.21
G-Re	7	469	276	1.3	2.0	2.6	4.3	3.2	4.1	1.2	1.0	−1.3	−2.3	0.58	0.31
Total	151	10117	4738	1.0	1.9	1.6	3.3	2.7	3.9	1.6	1.2	−0.7	−1.5	0.43	0.19

No. pix: number of pixels, Obs.: Number of observations, $\overline{qpe}$: mean rainfall in corresponding pixels, $\overline{gau}$: mean rainfall in corresponding raingauges, RMSD: Root mean square differenences rRMSD: relative RMSD, R: Pearson correlation coefficient. The * symbol indicates conditional statistics. S. Al: Basin located at South of rambla of El Albujón, Al: Rambla of El Albujón basin, NE. Al: Basin located at Northeast of rambla of El Albujón, G-No: Rambla of Nogalte basin, R.Pli: Pliego river basin and G-Re: Subasin of Guadalentín river where the Reguerón channel is located.

Table 2. Statistics of the multiplicative component of the differences between rain gauges and the corresponding pixels of PERSIANN-CCS and radar. Obs. stands for number of observations.

	Obs.	*Obs	Bias (dB)	Scatter (dB)
2009 event				
PERSIANN-CCS	4664	1200	−0.06	5.00
Radar CAPPI	4752	1563	−2.38	3.37
2012 event				
PERSIANN-CCS	2289	770	−0.19	7.06
2016 Event				
PERSIANN-CCS	9380	2375	−0.60	5.36
Radar SRI	10,117	5118	−2.41	3.25

3.2. Hourly Monitoring of Differences

The analysis of the temporal evolution of the intensities and of the agreement measures between the two estimates give an indication of the virulence of the storm as well as the possible hydrological response of the basin. This analysis is also important to validate precipitation products as it allows us to assess whether the analyzed QPE, in the absence of other sources of information, correctly estimates the evolution of the storm or the quantities precipitated in small basins prone to flash floods.

Figure 3a shows the evolution of the mean conditional precipitation according to rain gauges and the conditional MAD for the whole study area. Figure 3b–d shows the same no-conditional statistics.

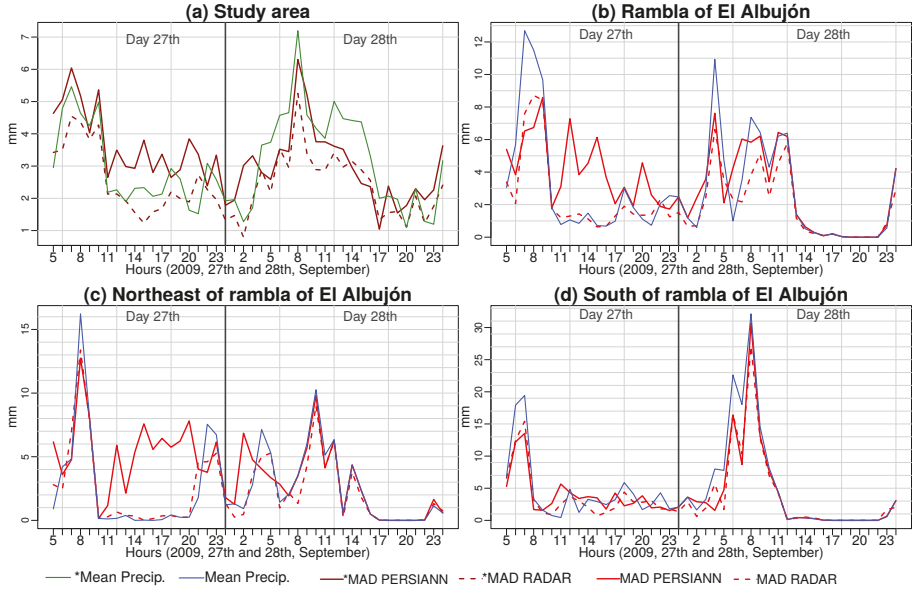


Figure 3. Temporal variation of Mean Absolute Differences (MAD) and rain gauge mean precipitation (Mean Precip.) in different spatial areas of the study area (PERSIANN-CCS and radar) during the 2009 event.

In the study area, both PERSIANN-CCS and radar *MAD values are very similar to $\overline{gau}$ when this value is high (Figure 3a). The *MAD of radar tends to be low when the $\overline{gau}$ is low, while this is not the case with PERSIANN-CCS. If we focus on the rain gauges of the Albujón basin we see that the MAD is proportional to the intensity of the precipitation and the peaks of MAD in both PERSIANN-CCS

and radar coincide with the two peaks of $\overline{gau}$. For the other two basins the results are similar. Both PERSIANN-CCS and radar have very high MAD for the magnitude of $\overline{gau}$.

Figure 4 shows the results for the 2012 event. MADs coincides with average precipitation at peak precipitation levels and again there is overestimation when average precipitation is low.

The results of the 2016 event are shown in Figure 5. In the study area, the temporal evolution of the *MAD indicates that this statistic, as in previous events, is closely related to the precipitation recorded by the rain gauges, although PERSIANN-CCS overestimates precipitation when it was low. The *MADs of PERSIANN-CCS shows very erratic estimations, with strong ups and downs in very short periods of time. In this sense, radar *MADs are much more constant and correlated for both high and low mean precipitation values. Radar patterns are more accurate than those of PERSIANN-CCS. The results for the three basins most affected by the storm (Figure 5b–d) point to a much more similar behaviour between the two QPEs. Again, average precipitation peaks reflect peaks in MAD in the two QPEs, although the MAD values of PERSIANN-CCS are usually higher.

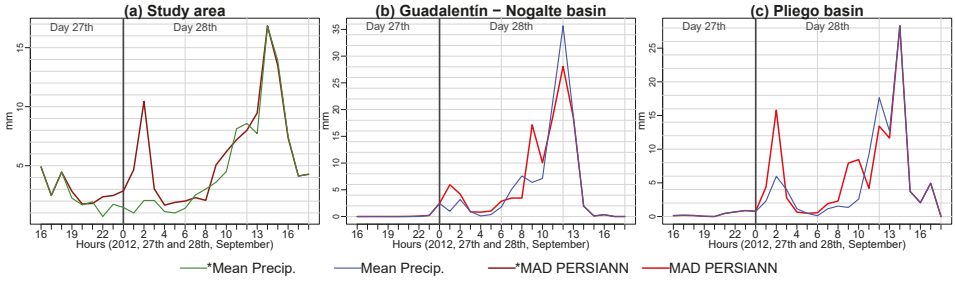


Figure 4. Temporal variation of Mean Absolute Differences (MAD) and rain gauge mean precipitation (Mean Precip.) in different spatial areas of the study area (PERSIANN-CCS) during the 2012 event.

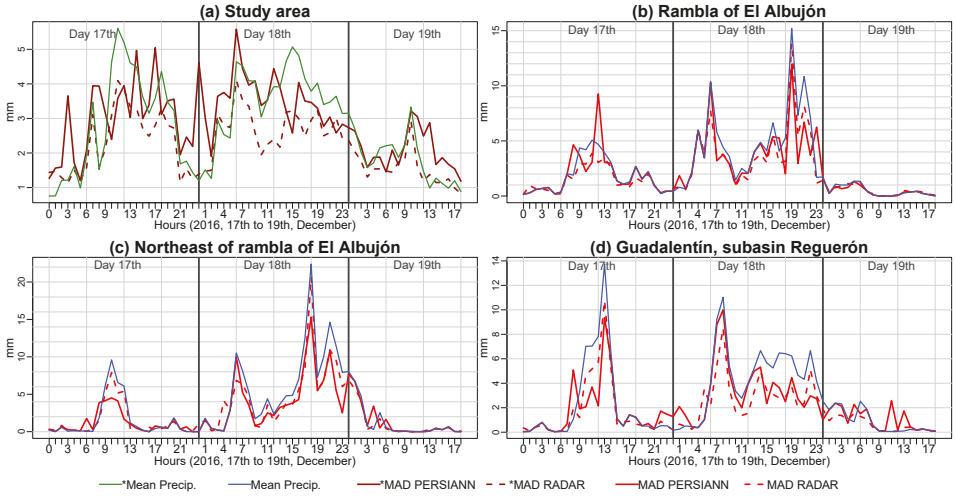


Figure 5. Temporal variation of Mean Absolute Differences (MAD) and rain gauge mean precipitation (Mean Precip.) in different spatial areas of the study area (PERSIANN-CCs and radar) during the 2016 event.

3.3. Side-By-Side Comparison of Accumulated Precipitation

To correctly validate QPEs, the map of precipitation accumulations of the whole event (Figure 6) was compared with the map of the differences observed (Figure 7) for the QPEs and the three events analyzed. Each map was generated with the native resolution of the QPE, which is evident from the appearance of the maps, especially in the case of the 2009 event, when the radar worked in short-range mode. To interpolate the rain gauge results (Figures 6c,f,i and 7) we used Ordinary kriging with automatic fitting procedures using the R geostatistical library *gstat* [37].

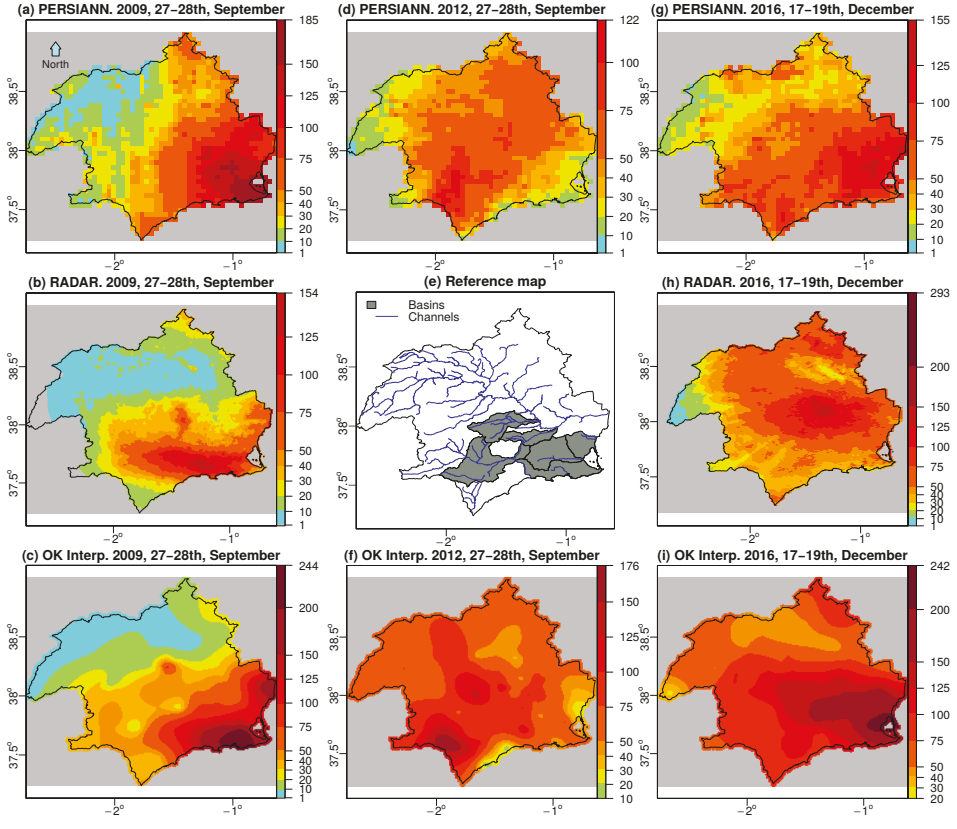


Figure 6. Total precipitation accumulations of the three events in mm. WGS projection (EPSG code 4326). Maps (**c**,**f**,**i**) are the result of an Ordinary kriging interpolation. Maps (**a**–**c**) from 0500 UTC 27 September to 0100 UTC 30 September 2009. Maps (**d**,**f**) from 1600 UTC 27 September to 1900 UTC 28 September 2012. (**g**–**i**) from 0000 UTC 17 December to 1900 UTC 19 December 2016.

The 2009 event shows an east-west rainfall gradient according to PERSIANN-CCS, with the highest values in the east. The maximum value of this QPE is 184 mm and the average value is 50 mm. The radar estimate for the 2009 episode, on the other hand, seems to be distributed in horizontal bands, with the highest values concentrated at the latitude of the coastline of the Region of Murcia descending towards the north and towards the south. The maximum radar accumulation value in this episode is 154 mm and the average is 30.7 mm, both much lower than the values provided by PERSIANN-CCS. With regard to the interpolation of differences (Figure 7a,b). The mean of the absolute values of the map is 17.7 mm, while the same statistic on the radar map is 19.4 mm. With respect to the maximums,

the PERSIANN-CCS is 104.5 mm and the radar is 169.2 mm. The map of interpolated radar differences (Figure 7b) shows very strong underestimations near Cartagena and the mouth of the Segura river, while these values tend to fall in a north-westerly direction. There is a small nucleus of overestimation around the city of Lorca.

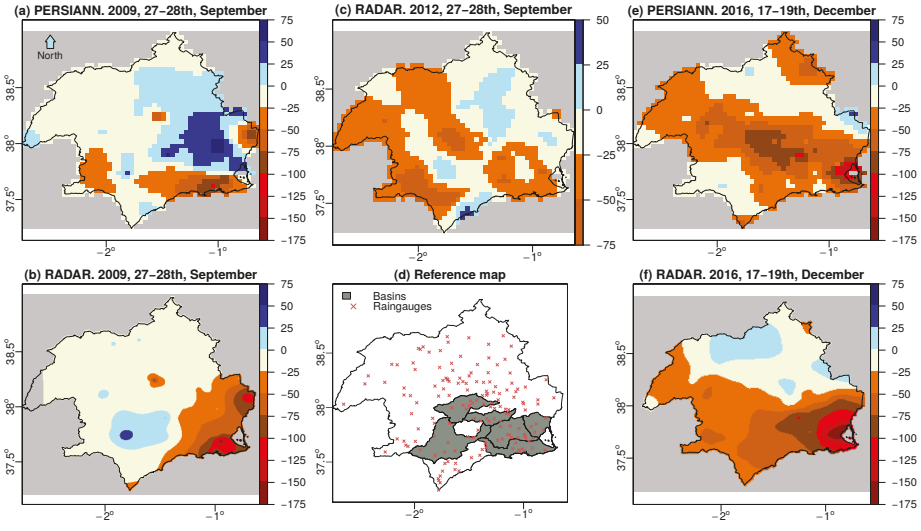


Figure 7. Interpolation by Ordinary kriging of the total accumulation differences (QPE pixel minus rain gauge) in each event in mm. WGS84 projection (EPSG code 4326). Maps (**a**,**b**) from 0500 UTC 27 September to 0100 UTC 30 September 2009. Maps (**c**) from 1600 UTC 27 September to 1900 UTC 28 September 2012. Maps (**e**,**f**) from 0000 UTC 17 December to 1900 UTC 19 December 2016.

With respect to the 2012 episode (Figure 6c), the largest accumulations estimated by PERSIANN-CCS are around the Nogalte and Guadalentín basins, where the largest accumulations actually occurred during this storm. The estimated accumulation values tend to fall towards the northwest and southeast. The interpolation of the differences identifies the areas where PERSIANN-CCS strongly underestimates the greatest rainfall accumulations recorded by the rain gauges.

Figures 6e and f show the precipitation accumulations of PERSIANN-CCS and radar for the 2016 event. Apart from the different spatial resolutions, both are similar except that PERSIANN-CCS places the greatest accumulations near the SE coast and radar places them in the centre of the study area. With regard to the differences, Figures 7e,f show that the greatest differences are found in the SE coast, although, due to its size, these differences are greater in the case of radar.

3.4. Spatial Statistics

Using the radar precipitation as baseline, spatial statistics were calculated for PERSIANN-CCS and radar. Figure 8 highlights these spatial relationships. The upscaling procedure applied for radar consisted in transferring values from the high-resolution raster cells to each one of the 0.04° grid cells using bilinear interpolation as implemented of the resample function of the raster R package [38]. Considering that the resampling was applied at an hourly timescale in which precipitation is assumed to be a smoothly varying variable within each 0.04° grid cell, we consider the bilinear interpolation to be a suitable technique with low impact on our results. The same procedure was applied by Zambrano-Bigiarini et al. (2017) [9] on a daily scale. In the 2009 event, the highest RMSD values appear in the coast, where the maximum rainfall values were recorded. RMSD decreases in a SE-NW direction. 83% of the study area shows RMSD equal to or smaller than 2 mm/h, so the agreement between both

QPEs is high. RMSD values for the 2016 event are not homogeneously distributed and the differences values are higher near the coast, although not exactly where rainfall was more intense. In this case, the RMSD values below 2 mm/h decrease to 56 percent, while 1.2% of the pixels has an RMSD equal to or greater than 4. The area covered by very low RMSD values of less than 1 mm/h is much smaller than in the 2009 episode.

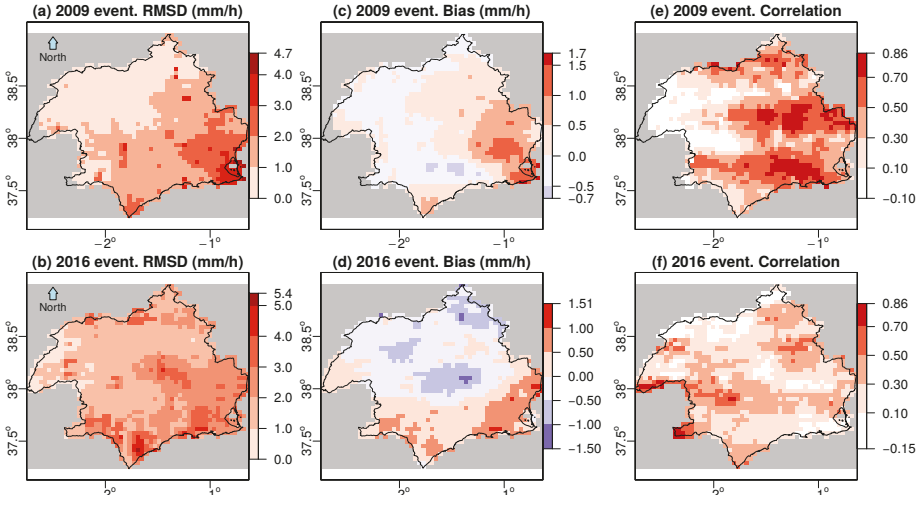


Figure 8. Comparision statistics between PERSIANN-CCS and radar hourly precipitation. WGS84 projection (EPSG code 4326). Maps (**a**,**c**,**e**) from 0500 UTC 27 September to 0100 UTC 30 September 2009. Maps (**b**,**d**,**f**) from 0000 UTC 17 December to 1900 UTC 19 December 2016.

Figure 8c,d shows the bias of each pixel. In the 2009 event, the pixels with a bias lower than 0 (assuming the radar data as the baseline) is 41%, which is much lower than for the 2016 event. As regards the correlation (Figure 8e,f), high correlation values are much more frequent in the 2009 event. Both QPEs are similar in the 2009 event and less so in the 2016 event.

3.5. Hyetographs

Finally, the precipitation of three pluviometers and the corresponding pixels of the QPEs were compared. Three rain gauges were selected for each event: the ones with the highest hourly rainfall intensity (Figures 9a,d and 10a), the ones with the highest rainfall accumulation (Figures 9b,e and 10b), and ones with the rainfall accumulation closest to the average cumulative precipitation of all the rain gauges (Figures 9c,f and 10c).

Starting with the precipitation observed in the 2009 hietographs, it is clear that both QPEs significantly underestimate rainfall. In the case of the rain gauge with the highest hourly intensity, both QPEs correctly detect the presence of precipitation, but not its quantity. In the hour with the highest intensity, radar is more accurate, in the hour with the second highest intensity, PERSIANN-CCS does not detect precipitation and radar shows minimum intensity. In the case of the highest accumulation rain gauge (Figure 9b), the highest precipitation intensities are not captured by any of the QPEs. The same can be said for the most representative total accumulation rain gauge. These results are similar for the 2012 (Figure 9d–f) and 2016 events (Figure 10).

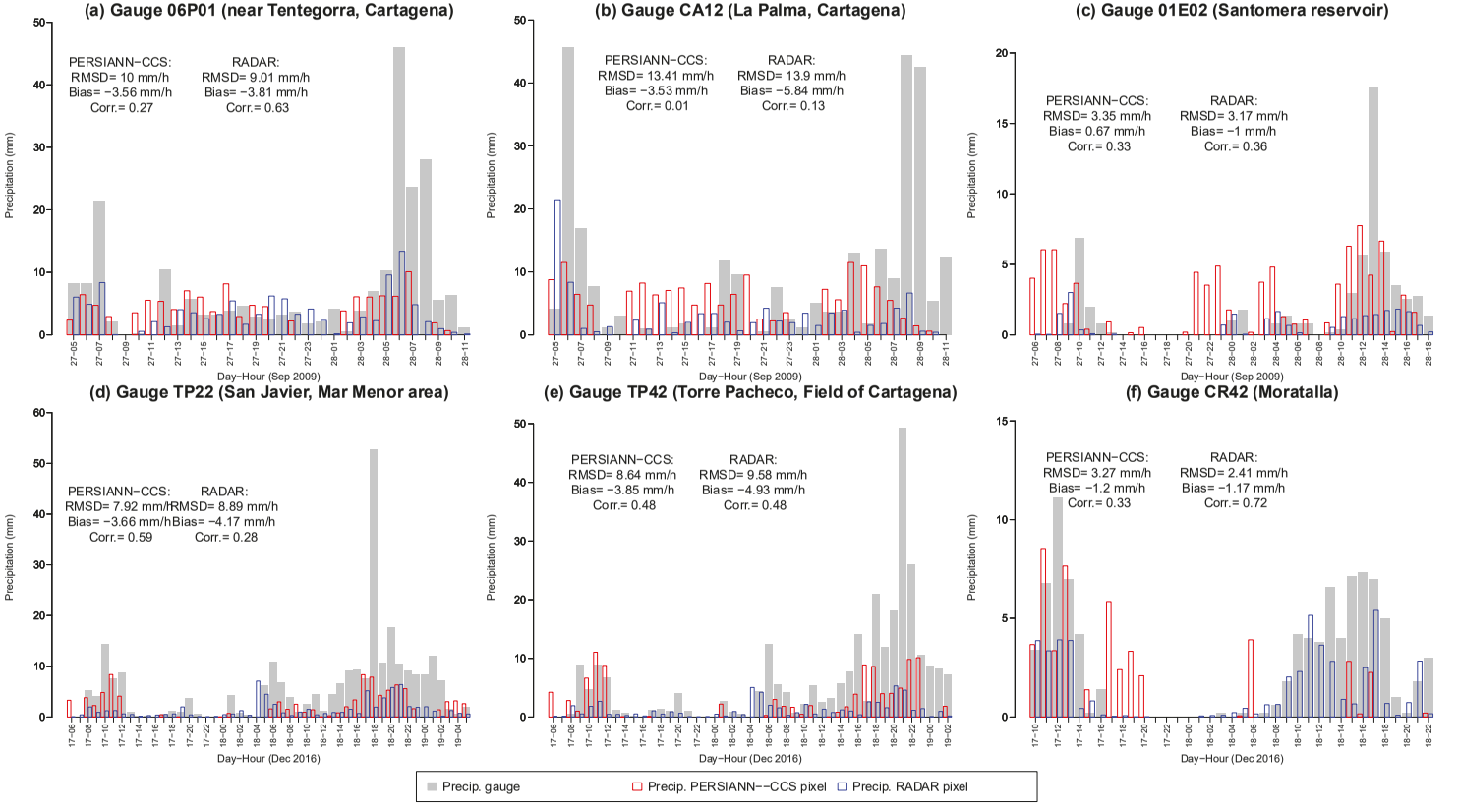


Figure 9. Hyetographs and associated statistics of PERSIANN-CCS, radar and the corresponding rain gauges of SIAM or SAIH-Segura. (**a–c**), 2009 event. (**d–f**), 2016 event.

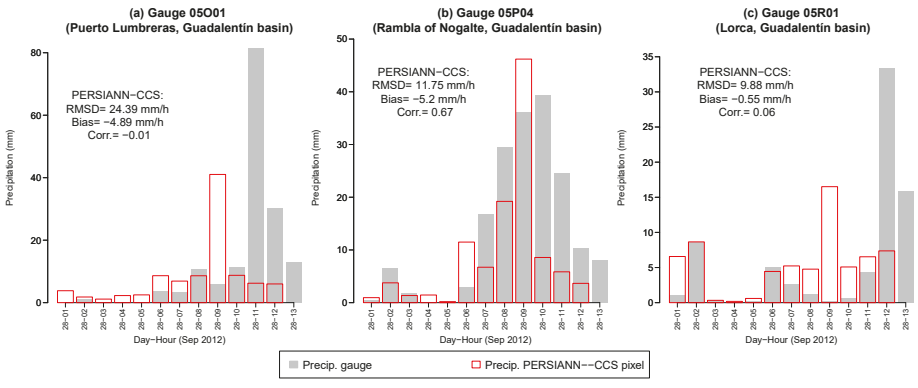


Figure 10. Hyetographs and associated statistics of PERSIANN-CCS and the corresponding rain gauges of SAIH-Segura.

4. Discussion and Conclusions

A comparison of three QPEs with two rain gauge networks has been carried out for the most severe rainfall events in the south-east of the Iberian Peninsula in the last decade. The objective was to know if either of the QPEs could be of use in monitoring such events. The results for the three events are similar: neither PERSIANN-CCS nor radar (both without empirical calibration from rain gauges) are acceptable QPEs for real-time monitoring. The larger the rainfall intensity, the greater the disagreement of the two QPEs with the rain gauges. When aggregated data are used (Table 1), the relative agreement statistics rRMSD and *rRMSD are higher than the mean precipitation. The underestimation is very high when rainfall recorded by rain gauges is also very high. The total cumulations of the QPEs also show significant differences. It does not seem that radar is much more accurate than PERSIANN-CCS, despite its larger spatial resolution and its commonly higher effectiveness. The bias in dB has high values when the radar is analyzed. According to our knowledge, this could be interpreted as the existence of a certain margin to improve radar QPEs with a simple global bias correction as explained in Germann et al. (2006) [36] for the Swiss radar network.

The results of this work agree with those of several previous contributions. On a daily scale, Burcea et al. (2012) [35] also found that meteorological radar in the Moldavian Plaeau underestimated rainfall recorded by rain gauges. Changing the time scale from daily to ten minutes, similar results have also been documented in Seoul (South Corea) [39].

Several studies have documented differences in satellite-based QPE estimations when precipitation intensities are high. On a daily scale, but for a larger spatial scope, such as Chile, Zambrano-Bigiarini et al. (2017) [9] pointed out that PERSIANN-CCS and other satellite-based QPEs were able to correctly identify the occurence of no-rain events, but had low accuracy when classifying precipitation intensities during rainy days. Similar patterns have been identified in other mountain areas using the TRMM (Tropical Rainfall Measuring Mission) Multi-Satellite Precipitation Analysis (TMPA) 3B42b6 [40]. In Europe, other satellite-based QPEs have been evaluated, such as the CMORPH (Climate Center Morphing technique) of NOAA,a satellite QPE similar to PERSIANN-CCS, with a spatial resolution of 8×8 km^2 and a temporal resolution of 30 min [41]. In this case, the data were temporarily aggregated to one hour and five heavy rainfall events were analyzed in three European mountain areas located in the Italian Alps and the Massif Central mountain range. Evaluation of the rainfall estimates, unlike in this study, was based on high-quality rain gauge calibrated radar rainfall fields. These authors also highlight that this QPE, without a empirical calibration, underestimates rainfall when analysing heavy precipitation events. Nikolopoulos et al. (2013) [5] reported similar results for TRMM TMPA, CMORP and PERSIANN-CCS, pointing to the problems of underestimation

of flash-floods if such products are used as input for hydrological modelling in mountainous areas. Chen et al. (2013) [42] compared four satellite-based QPEs, including PERSIANN-CCS, and radar, for an extreme event, the Moratok typhoon over Taiwan. This event was larger (more than two days) and more extreme (2777 mm with a fairly constant rainfall intensity) than the events studied in this paper. However, and despite the differences in meteorological causes and in the structure of the events, the results with PERSIANN-CCS are very similar to those presented in this paper.

Although our results are based on only three events, and do not provide statistical significance, they represent an example of satellite-based and ground-based QPE accuracies in a Mediterranean basin for three extreme storm events that caused major floods, and point to the severe underestimation shown by the QPEs in all the events. We only analysed convective events; it would be interesting to run a comparative analysis distinguishing convective and stratiform events. Recently, other authors [22] evaluated the performance of PERSIANN family product to reproduce daily rainfall for the period 2003–2015 at global scale. They concluded that the better performance of PERSIANN-CDR in comparison to PERSIANN-CCS, is justified by the bias adjustment of PERSIANN-CDR on a monthly scale using ground observations (Global Precipitation Climatology Project, GPCP data). From the evaluation at global scale, according to Nguyen et al. (2018) [22], PERSIANN-CCS estimates higher rainfall over continents except for Europe. These results confirm the underestimations identified over South East of Spain from the present work.

This work is based on a network of rain gauges whose information is assumed to be ground truth and that are not randomly or regularly distributed in the territory. Due to the specific purposes of the two rain gauges networks, irrigated crop areas, valley bottoms, coastal plains and places where there are relevant hydraulic infrastructures for flood management or water accumulation, mainly reservoirs, are over-represented. On the other hand, forest or scrub areas, the upper part of the basin and the mountain peaks are under-represented (Figure 1). It is not clear the extent to which this over-representation can influence the bias of the statistics used in the research, and we suggest that this is an interesting topic to be tackled with data from other more randomly located rain gauge networks.

According to the results obtained in this work and in agreement with the literature, all these precipitation products still present serious problems when it comes to quantitatively estimating rainfall during very heavy precipitation events. Both PERSIANN-CCS and other satellite-based QPEs present the common problem of underestimating high precipitation intensities. However, it should be noted that due to its close to global coverage, its high spatial resolution, its high temporal resolution and its short lag time, this satellite-based QPE presents very good characteristics for a local calibration of an empirical type based on a network of rain gauges located in the field and providing data in real time. In case of rain gauge failures, the applicability and availability of precipitation data obtained from satellite sources, such as PERSIANN-CCS, could be of value since the methods they use to collect information are independent of local conditions.

The three analysed QPEs do not reproduce the spatio-temporal variability of heavy rainfall events. However, it is possible that they could serve as predictors when interpolating rainfall on sub-daily scales using machine learning regression algorithms. Such non-parametric algorithms outperform parametric linear regression or GLM if the data have a significant proportion of noise or if the assumptions of the linear models are not met. In particular, two of these algorithms, random forest [43] and neural networks [44], have been used to interpolate air temperature using land surface temperature retrieved from remote sensing imagery) as a predictor.

Author Contributions: The three authors contributed equally to this work.

Funding: This work is the result of a postdoctoral contract funded by Saavedra Fajardo programme (Ref. 20023/SF/16) of the Consejería de Educación y Universidades of CARM (Autonomous Community of Murcia Region), by the Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia.

Acknowledgments: The support and availability of information from the Center for Hydrometeorology and Remote Sensing of University of California-Irvine (USA), from the Segura Basin Hydrological Confederation (CHS) and from Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario (IMIDA) of CARM

are acknowledged. Source radar data were provided by the Spanish Meteorological Agency (AEMET) of the Spanish Ministry of Farming and Fishing, Fooding and Environment. The authors thank Luis Bañón Peregrín, Juan Manzano Cano and José Miguel Gutiérrez (AEMET) for assistance in understanding the radar-rainfall data products available. We also thank the four anonymous reviewers whose suggestions have substantially improved this manuscript.

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

References

1. 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*, 1834–1852. [CrossRef]
2. Sun, Q.; Miao, C.; Duan, Q.; Sorooshian, S.; Hsu, K.L. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Rev. Geophys.* **2018**, *56*, 79–107. [CrossRef]
3. Sorooshian, S.; Nguyen, P.; Sellars, S.; Braithwaite, D.; AghaKouchak, A.; Hsu, K. Satellite-based remote sensing estimation of precipitation for early warning systems. In *Extreme Natural Hazards, Disaster Risks and Societal Implications*; Cambridge University Press: Cambridge, UK, 2014; pp. 99–112.
4. Bendix, J.; Fries, A.; Zárate, J.; Trachte, K.; Rollenbeck, R.; Pucha-Cofrep, F.; Paladines, R.; Palacios, I.; Orellana, J.; Oñate Valdivieso, F.; et al. RadarNet-Sur First Weather Radar Network in Tropical High Mountains. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1235–1254. [CrossRef]
5. Nikolopoulos, E.I.; Anagnostou, E.N.; Borga, M. Using High-Resolution Satellite Rainfall Products to Simulate a Major Flash Flood Event in Northern Italy. *J. Hydrometeorol.* **2013**, *14*, 171–185. [CrossRef]
6. Miao, C.; Ashouri, H.; Hsu, K.L.; Sorooshian, S.; Duan, Q. Evaluation of the PERSIANN-CDR Daily Rainfall Estimates in Capturing the Behavior of Extreme Precipitation Events over China. *J. Hydrometeorol.* **2015**, *16*, 1387–1396. [CrossRef]
7. Schiemann, R.; Erdin, R.; Willi, M.; Frei, C.; Berenguer, M.; Sempere-Torres, D. Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1515–1536. [CrossRef]
8. Ballari, D.; Castro, E.; Campozano, L. Validation of Satellite Precipitation (TRMM 3B43) in Ecuadorian Coastal Plains, Andean Higlands and Amazonian Rainforest. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; Copernicus GmbH: Prague, Czech Republic, 2016; Volume XLI-B8, pp. 305–311.
9. Zambrano-Bigiarini, M.; Nauditt, A.; Birkel, C.; Verbist, K.; Ribbe, L. Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 1295–1320. [CrossRef]
10. Barredo, J.I. Major flood disasters in Europe: 1950–2005. *Nat. Hazards* **2007**, *42*, 125–148. [CrossRef]
11. López-Martínez, F.; Gil-Guirado, S.; Pérez-Morales, A. Who can you trust? Implications of institutional vulnerability in flood exposure along the Spanish Mediterranean coast. *Environ. Sci. Policy* **2017**, *76*, 29–39. [CrossRef]
12. Serrano-Notivoli, R.; Martín-Vide, J.; Saz, M.A.; Longares, L.A.; Beguería, S.; Sarricolea, P.; Meseguer-Ruiz, O.; de Luis, M. Spatio-temporal variability of daily precipitation concentration in Spain based on a high-resolution gridded data set. *Int. J. Climatol.* **2017**, *38*, e518–e530. [CrossRef]
13. López-Bermúdez, F.; Conesa-García, C.; Alonso-Sarría, F. Floods: Magnitude and Frequency in Ephemeral Streams of the Spanish Mediterranean Region. In *Dryland Rivers: Hydrology and Geomorphology of Semi-Arid Channels*; John Wiley & Sons: Hoboken, NJ, USA, 2002; pp. 329–350.
14. 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.* **2015**, *116*. [CrossRef]
15. 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. JHM574.1. [CrossRef]
16. Pellicer-Martínez, F.; Martínez-Paz, J.M. Probabilistic evaluation of the water footprint of a river basin: Accounting method and case study in the Segura River Basin, Spain. *Sci. Total Environ.* **2018**, *627*, 28–38. [CrossRef] [PubMed]

17. Gomariz-Castillo, F.; Alonso-Sarría, F.; Cabezas-Calvo-Rubio, F. Calibration and spatial modelling of daily ET_0 in semiarid areas using Hargreaves equation. In *Earth Science Informatics*; Springer: Berlin/Heidelberg, Germany, 2017.
18. Pellicer-Martínez, F.; Martínez-Paz, J.M. Grey water footprint assessment at the river basin level: Accounting method and case study in the Segura River Basin, Spain. *Ecol. Indic.* **2016**, *60*, 1173–1183. [CrossRef]
19. Giordano, R.; Pagano, A.; Pluchinotta, I.; Olivo del Amo, R.; Hernandez, S.M.; Lafuente, E.S. Modelling the complexity of the network of interactions in flood emergency management: The Lorca flash flood case. *Environ. Model. Softw.* **2017**, *95*, 180–195. [CrossRef]
20. García-Ayllón, S. GIS Assessment of Mass Tourism Anthropization in Sensitive Coastal Environments: Application to a Case Study in the Mar Menor Area. *Sustainability* **2018**, *10*, 1344. [CrossRef]
21. Ministerio de Fomento. Dirección General de Carreteras. In *Máximas Lluvias Diarias en la España Peninsular*; Ministerio de Fomento: Madrid, Spain, 1999.
22. 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. Discuss.* **2018**, in review. [CrossRef]
23. Nguyen, P.; Sellars, S.; Thorstensen, A.; Tao, Y.; Ashouri, H.; Braithwaite, D.; Hsu, K.; Sorooshian, S. Satellites Track Precipitation of Super Typhoon Haiyan. *Eos Trans. Am. Geophys. Union* **2014**, *95*, 133–155. [CrossRef]
24. Nguyen, P.; Thorstensen, A.; Sorooshian, S.; Hsu, K.; AghaKouchak, A. Flood Forecasting and Inundation Mapping Using HiResFlood-UCI and Near-Real-Time Satellite Precipitation Data: The 2008 Iowa Flood. *J. Hydrometeorol.* **2015**, *16*, 1171–1183. [CrossRef]
25. Karbalaee, N.; Hsu, K.; Sorooshian, S.; Braithwaite, D. Bias adjustment of infrared-based rainfall estimation using Passive Microwave satellite rainfall data. *J. Geophys. Res. Atmos.* **2017**, *122*, 3859–3876. [CrossRef]
26. Fukao, S.; Hamazu, K.; Doviak, R.J. *Radar for Meteorological and Atmospheric Observations*; Springer: Tokyo, Japan, 2014.
27. Velasco-Forero, C.A.; Sempere-Torres, D.; Cassiraga, E.F.; Gómez-Hernández, J.J. A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Adv. Water Resour.* **2009**, *32*, 986–1002. [CrossRef]
28. 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]
29. 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]
30. Hill, D.J.; Baron, J. radar.IRIS: A free, open and transparent R library for processing Canada's weather radar data. *Can. Water Resour. J.* **2015**, *40*, 409–422. [CrossRef]
31. Zawadzki, I. On Radar-Raingage Comparision. *J. Appl. Meteorol.* **1975**, *14*, 1430–1436. [CrossRef]
32. Freedman, D.; Pisani, R.; Purves, R. *Statistics*, 4 ed.; Viva Books: New Delhi, India, 2009.
33. Germann, U.; Berenguer, M.; Sempere-Torres, D.; Zappa, M. REAL—Ensemble radar precipitation estimation for hydrology in a mountainous region. *Q. J. R. Meteorol. Soc.* **2009**, *135*, 445–456. [CrossRef]
34. Speirs, P.; Gabella, M.; Berne, A. A Comparison between the GPM Dual-Frequency Precipitation Radar and Ground-Based Radar Precipitation Rate Estimates in the Swiss Alps and Plateau. *J. Hydrometeorol.* **2017**, *18*, 1247–1269. [CrossRef]
35. Burcea, S.; Cheval, S.; Dumitrescu, A.; Antonescu, B.; Bell, A.; Breza, T. Comparision Between Radar Estimated and Rain Gauge Measured Precipitation in the Moldavian Plateau. *Environ. Eng. Manag. J.* **2012**, *11*, 723–731.
36. Germann, U.; Galli, G.; Boscacci, M.; Bolliger, M. Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.* **2006**, *132*, 1669–1692. [CrossRef]
37. Pebesma, E.J. Multivariable geostatistics in S: the gstat package. *Comput. Geosci.* **2004**, *30*, 683–691. [CrossRef]
38. Hijmans, R.J. raster: Geographic Data Analysis and Modeling. 2016. Available online: https://cran.r-project.org/web/packages/raster/ (accessed on 10 May 2018).
39. Yoon, S.-S.; Lee, B. Effects of Using High-Density Rain Gauge Networks and Weather Radar Data on Urban Hydrological Analyses. *Water* **2017**, *9*, 931. [CrossRef]

40. Scheel, M.L.M.; Rohrer, M.; Huggel, C.; Santos Villar, D.; Silvestre, E.; Huffman, G.J. Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2649–2663. [CrossRef]
41. Zhang, X.; Anagnostou, E.N.; Frediani, M. Using NWP Simulations in Satellite Rainfall Estimation of Heavy Precipitation Events over Mountainous Areas. *J. Hydrometeorol.* **2013**, *14*, 1844–1858. [CrossRef]
42. Chen, S.; Hong, Y.; Cao, Q.; Kirstetter, P.E.; Gourley, J.J.; Qi, Y.; Zhang, J.; Howard, K.; Hu, J.; Wang, J. Performance evaluation of radar and satellite rainfalls for Typhoon Morakot over Taiwan: Are remote-sensing products ready for gauge denial scenario of extreme events? *J. Hydrol.* **2013**, *506*, 4–13. [CrossRef]
43. Xu, Y.; Knudby, A.; Ho, H.C. Estimating daily maximum air temperature from MODIS in British Columbia, Canada. *Int. J. Remote Sens.* **2014**, *35*, 8108–8121. [CrossRef]
44. Jang, J.; Viau, A.; Anctil, F. Neural network estimation of air temperatures from AVHRR data. *Int. J. Remote Sens.* **2004**, *25*, 4541–4554. [CrossRef]

remote sensing

MDPI

Technical Note

Downscaling of Satellite OPEMW Surface Rain Intensity Data

Angela Cersosimo [1,*], Salvatore Larosa [1], Filomena Romano [1], Domenico Cimini [1,2], Francesco Di Paola [1], Donatello Gallucci [1], Sabrina Gentile [1,2], Edoardo Geraldi [1,3], Saverio Teodosio Nilo [1], Elisabetta Ricciardelli [1], Ermann Ripepi [1] and Mariassunta Viggiano [1]

[1] Institute of Methodologies for Environmental Analysis, National Research Council (IMAA-CNR), 85100 Potenza, Italy; salvatore.larosa@imaa.cnr.it (S.L.); filomena.romano@imaa.cnr.it (F.R.); domenico.cimini@imaa.cnr.it (D.C.); francesco.dipaola@imaa.cnr.it (F.D.P.); donatello.gallucci@imaa.cnr.it (D.G.); sabrina.gentile@imaa.cnr.it (S.G.); edoardo.geraldi@imaa.cnr.it (E.G.); saverio.nilo@imaa.cnr.it (S.T.N.); elisabetta.ricciardelli@imaa.cnr.it (E.R.); ermann.ripepi@imaa.cnr.it (E.R.); mariassunta.viggiano@imaa.cnr.it (M.V.)

[2] Center of Excellence Telesensing of Environment and Model Prediction of Severe events (CETEMPS), University of L'Aquila, 67100 L'Aquila, Italy

[3] Institute for Archaeological and Monumental Heritage, National Research Council (IBAM-CNR), 85100 Potenza, Italy

* Correspondence: angela.cersosimo@imaa.cnr.it; Tel.: +39-0971-427500

Received: 9 October 2018; Accepted: 5 November 2018; Published: 8 November 2018

Abstract: This paper presents a geostatistical downscaling procedure to improve the spatial resolution of precipitation data. The kriging method with external drift has been applied to surface rain intensity (SRI) data obtained through the Operative Precipitation Estimation at Microwave Frequencies (OPEMW), which is an algorithm for rain rate retrieval based on Advanced Microwave Sounding Units (AMSU) and Microwave Humidity Sounder (MHS) observations. SRI data have been downscaled from coarse initial resolution of AMSU-B/MHS radiometers to the fine resolution of Spinning Enhanced Visible and InfraRed Imager (SEVIRI) flying on board the Meteosat Second Generation (MSG) satellite. Orographic variables, such as slope, aspect and elevation, are used as auxiliary data in kriging with external drift, together with observations from Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager (MSG-SEVIRI) in the water vapor band (6.2 μm and 7.3 μm) and in thermal-infrared (10.8 μm and 8.7 μm). The validation is performed against measurements from a network of ground-based rain gauges in Southern Italy. It is shown that the approach provides higher accuracy with respect to ordinary kriging, given a choice of auxiliary variables that depends on precipitation type, here classified as convective or stratiform. Mean values of correlation (0.52), bias (0.91 mm/h) and root mean square error (2.38 mm/h) demonstrate an improvement by +13%, −37%, and −8%, respectively, for estimates derived by kriging with external drift with respect to the ordinary kriging.

Keywords: surface rain intensity; kriging with external drift; PEMW; MSG; SEVIRI; downscaling

1. Introduction

Rainfall is of primary importance in many scientific fields, such as meteorology, hydrology, agriculture, ecology and other environmental sciences [1]. Precipitation intensity can be estimated with different techniques, including rain gauge, ground-based radar, and satellite remote sensing observations. Estimates from satellite are particularly relevant, as they assure a global coverage of the Earth [2]. The major drawback of satellite rainfall remote sensing lies in their coarse spatial resolution, which hinders the investigation of spatial variability. This calls for the development of techniques to improve the spatial resolution of satellite rainfall remote sensing data.

A large number of downscaling methods has been developed and applied in the last few years. Most of these methods are based on the correlation between rainfall and environmental information such as latitude, longitude, altitude, slope, aspect and other orographic characteristics. In particular, Next Generation Radar (NEXRAD) daily precipitation fields were downscaled from 16 km to 4 km by considering orographic effects on precipitation distribution [3]. This method consists of three parts, namely the rain-pixel clustering, the multivariate regression and the random cascade. In [4], a statistical downscaling method—based on the relationships between precipitation and both terrain factors (e.g., slope, aspect and roughness) and meteorological conditions (e.g., humidity and temperature)—was proposed to disaggregate the Tropical Rainfall Measuring Mission (TRMM) 3B42 from 25 km to 1 km. Other studies introduced the positive relation between vegetation and precipitation through the use of the Normalized Difference Vegetation Index (NDVI). Among these, Ref. [5] explored the relation between TRMM rainfall estimates and NDVI at different spatial scales; the derived relation has then been used to develop a downscaling method based on an exponential regression model. In [6], a multiple linear regression model was developed using both NDVI and Digital Elevation Model (DEM) as independent variables. Ref. [7] presented a geostatistical downscaling procedure, namely a geographically-weighted regression kriging, based on the relationship between precipitation and other variables, such as NDVI and DEM, to downscale the TRMM 3B43 product from 25 km to 1 km. Two methods were used in [7]—namely a geographical difference analysis and a geographical ratio analysis—to calibrate the downscaled TRMM precipitation data. Ref. [8] applied an integrated downscaling method, based on environmental information such as vegetation, topography, drought and albedo derived from Moderate Resolution Imaging Spectroradiometer (MODIS) products. Two different downscaling approaches—a multiple linear regression and an artificial neural network—were compared and used to downscale TRMM precipitation data from 25 km to 1 km spatial resolution. The above studies demonstrate that downscaled precipitation data better capture the spatial variability compared to the original datasets.

In this work, we downscale precipitation data derived from the operational version of the Precipitation Estimation at Microwave Frequencies (PEMW), an algorithm for surface rain intensity (SRI) retrievals, developed at the Institute of Methodologies for Environmental Analysis of the National Research Council of Italy (CNR-IMAA) [9]. SRI data by PEMW is derived from Advanced Microwave Sounding Units (AMSU) and Microwave Humidity Sounder (MHS) radiometers on board Low-Earth-Orbit satellites and their spatial resolution range from 16 km at nadir to 51 km at maximum scanning angle. The Operative version of PEMW (OPEMW) was validated against simultaneous ground-based observations from weather radar systems and rain gauges [10]. OPEMW SRI data are downscaled from the AMSU-B/MHS spatial resolution (16 km at the sub satellite point—SSP) to the finer MSG-SEVIRI spatial resolution (3 km at the SSP) to possibly improve their accuracy and capture higher-resolution spatial variability. Among the methods mentioned above, a kriging technique is chosen, namely kriging with external drift, as it was shown to provide estimations and associated errors with satisfactory results even in the presence of few input data. Kriging with external drift different variables and their combinations are used as auxiliary data to improve the disaggregation process. Particularly, the Brightness Temperatures (BTs) from the MSG-SEVIRI [11] are exploited together with some environmental information, such as elevation, slope and aspect. The MSG-SEVIRI observations used here are from four channels, two centered in water vapour bands (at 6.2 μm and 7.3 μm) and two in the thermal infrared (at 10.8 μm and 8.7 μm). The precipitation episodes are divided into two types (convective and stratiform), and the trend producing the best results are searched.

The paper is organized as follows: Section 2 provides a description of the data used for the implementation and the validation of the downscaling procedure; it also describes the applied methodology, focusing on the construction of the variogram and on the choice of auxiliary variables. Section 3 shows the validation comparing downscaled surface rain rate and rain gauge measurements. In Section 4, we draw our conclusions.

2. Materials and Methods

2.1. Data Set

This section describes the OPEMW. Furthermore, we analyze the data used as auxiliary variables in kriging with external drift, such as the MSG/SEVIRI channels and the orographic variables (slope, aspect and elevation). Finally, we discuss details of the validation dataset, obtained by a network of ground-based rain gauges.

2.1.1. OPEMW

OPEMW is the operational version of PEMW running at IMAA-CNR since 2010. A detailed description of PEMW software can be found in [9]; here, we only provide a brief discussion. PEMW consists of a rainfall estimation algorithm that exploits both the radiometric observations made at window channels of 89 GHz and 150 GHz and at the water vapor band of 183 GHz [9]. It exploits the window channels to detect the size of the precipitating particles and to sense low-level precipitation, while the three water vapor bands are used to discriminate precipitation at different altitudes (high convective system, middle-altitude stratiform precipitation) [10]. PEMW is based on the observations acquired from AMSU-B on board the National Oceanic and Atmospheric Administration's (NOAA) Polar Operational Environmental Satellites (POES), from MHS on board the European Polar System (EPS) and finally from the most recent NOAA POES. AMSU-B and MHS are respectively cross-track, line scanning microwave radiometers measuring radiances in five channels in the frequencies ranging from 89 GHz to 190 GHz. In particular, AMSU-B exploits channels at central frequencies of 89 GHz, 150 GHz, 183 $\pm$ 1 GHz, 183 $\pm$ 3 GHz and 183 $\pm$ 7 GHz, while MHS at central frequencies of 89 GHz, 157 GHz, 183 $\pm$ 1 GHz, 183 $\pm$ 3 GHz and 190 GHz. AMSU-B and MHS fly at a nominal altitude of 850 km, and they scan the Earth about $\pm 50^{\circ}$ from subsatellite point. Each channel has an antenna beam width of 1.1°. This provides a resolution of 16 km at nadir and 90 consecutive fields of view (FOVs) per scan. One of the main advantages related to the use of AMSU-B and MHS observations is the good spatial resolution, which is often an issue for MW instruments. The OPEMW algorithm has been validated against ground-based observations from a network of 20 weather radar systems and a network of more than 3000 rain gauges distributed over the Italian territory [10]. The data set used for the validation spans over one year of surface rain intensity (July 2011–June 2012). The validation shows 98% accuracy in correctly identifying rainy and non-rainy areas. The correlation coefficient is larger than 0.8 and 0.9 with respect to rain gauge and weather radars, respectively, though a binned analysis in the 0–15 mm h^{-1} range suggests that the algorithm tends to overestimate rain rate values below 6–7 mm h^{-1} and underestimate those above 6–7 mm h^{-1}.

2.1.2. MSG-SEVIRI

SEVIRI is the visible/infrared imager on board the geostationary MSG satellite. SEVIRI is characterized by high temporal (15 min) and spatial resolutions. SEVIRI is a 50 cm diameter aperture, line by line scanning radiometer, which has the capability to observe the Earth in 12 spectral channels [11]. Out of these, eleven channels, namely the Visible (VIS), Near-InfraRed (NIR) and InfraRed (IR) ones, cover the full disk and have an imaging sampling distance of 3 km at SSP. Conversely, the remaining twelfth channel, i.e., the High Resolution Visible (HRV) channel, covers half the full disk with a 1 km at SSP imaging sampling distance at subsatellite point [12,13].

The MSG-SEVIRI BTs from the water vapor (6.2 and 7.3 μm) and the thermal-IR (10.8 and 8.7 μm) channels are exploited as auxiliary variables in the implementation of the kriging with external drift method. We consider: (i) BT differences between thermal-IR at 10.8 μm and water vapor at 6.2 μm channels, as this information is useful to identify deep-convection areas [14–17]; (ii) BT differences between water vapor at 6.2 and 7.3 μm channels, to recognize areas of intense precipitation [18]; and finally (iii) BT difference between 8.7 μm and 10.8 μm (thermal-infrared) channels, useful to

discriminate liquid and ice cloud that could be associated to mid-level stratiform cloud (nimbostratus) and convective clouds, respectively.

2.1.3. Rain Gauge Network

The downscaled surface rain intensity data are validated against ground observations of rain rate measured by rain gauges distributed on the area of Basilicata region, Southern Italy (Figure 1). The considered rain gauge network is currently managed and maintained by the Functional Center Decentralized (FCD) of Basilicata [19]. The FCD carries out daily hydrometeorological monitoring and forecasting activities using meteorological data from national forecasting models and data recorded by ground stations. The FCD of Basilicata currently manages 63 monitoring stations homogeneously distributed on the regional territory. The sensors installed on each station vary depending on the monitored site and, generally, consist of rain gauges, hydrometers, thermometers, anemometers, barometers and radiometers. The data flow into an archive managed by the FCD and they are processed after quality control operations. The acquisition and processing of data measured by rain gauges are performed at different temporal intervals, ranging from 5 to 60 min. We consider only 44 rain gauges for the validation of the downscaling procedure, as these provide data at the shortest time interval (15 min). The rain rate is measured by double-tipping rain gauges and the unit of measurement of the water deposited on the ground is millimeters (mm). The rain gauges used in the validation are homogeneously distributed on the territory under analysis, as shown in Figure 1. More details on considered rain gauges are shown in Appendix B.

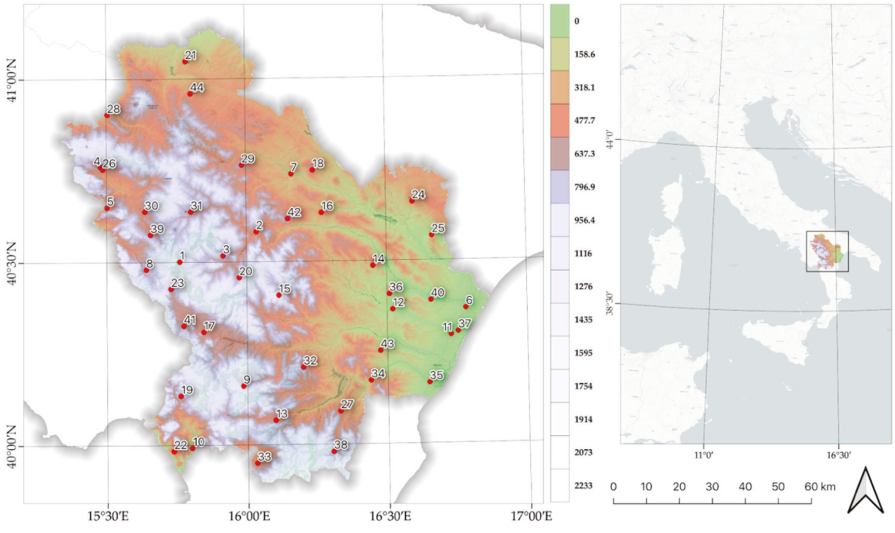


Figure 1. Geographical location of the used meteorological ground observation stations. In the right panel, blue is for water and white is for land.

2.1.4. Elevation, Slope, Aspect

Digital Terrain Models (DTMs) are a primary input to any modelling or quantification process involving the earth's topography. DTMs are raster files containing elevation data for each raster cell. DTMs are widespread for calculations, manipulations and further analysis of a geographic area, and are mainly used for extrapolating elevation information and other derived data, such as slope and aspect, related to the morphological characteristics of the study area. DTM datasets are widely used in GIS applications, and several built-in tools are available to turn the DTM into a derivative

map. There are several free DTM products available, which feature high accuracy and data resolution. The accuracy of the elevation is highly correlated to orographic features and DTMs play an important role in the orographic analysis. In this work, we used a DTM from Shuttle Radar Topography Mission (SRTM) with a spatial resolution of 3 arc seconds (≈90 m) in order to assess properly the quality of our algorithm. SRTM data were collected during the 11-day mission in February 2000. Since then, SRTM data were described in detail [20–22] and became accessible online for free [23] SRTM3 has 90 m spatial resolution at the equator and is provided in mosaiced 5 deg × 5 deg tiles for easy download and use. All tiles are produced from a seamless dataset to allow easy mosaicking. These tiles are available in both ArcInfo ASCII and GeoTiff, formatted to ease their use in a variety of image processing and GIS applications.

The slope map was obtained using the maximum gradient of the plane tangent to the surface of the ground at a certain point. One of the methods for calculating the slope in one point is the finite differences [24], for which the following calculation example is given. The east–west and north–south gradient for the center cell in the 3 × 3 floating window (kernel) is defined as follows:

$$\left[\frac{\partial z}{\partial y}\right] = \left[\frac{\left(z_{i-1,j-1} + 2 \times z_{i,j-1} + z_{i+1,j-1}\right) - \left(z_{i-1,j+1} + 2 \times z_{i,j+1} + z_{i+1,j+1}\right)}{8 \times \Delta y}\right], \quad (1)$$

$$\left[\frac{\partial z}{\partial x}\right] = \left[\frac{\left(z_{i+1,j+1} + 2 \times z_{i+1,j} + z_{i+1,j-1}\right) - \left(z_{i-1,j+1} + 2 \times z_{i-1,j} + z_{i-1,j-1}\right)}{8 \times \Delta x}\right], \quad (2)$$

where z represents the elevation of the eight surrounding cells and Δx, Δy specify the cell dimensions in the horizontal and lateral directions, respectively (Figure 2).

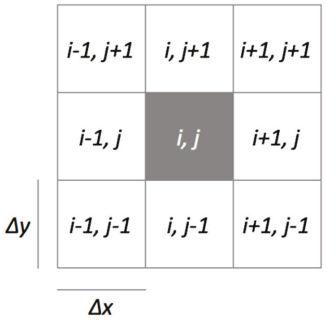


Figure 2. 3 × 3 kernel for slope calculation and i, j is the central cell. Δx and Δy are the dimensions of individual cells.

The slope angle is calculated with respect to the central cell i, j of the kernel window and is given by the surface gradient module:

$$\text{slope} = \frac{180}{\pi} \times \arctan\left(\sqrt{\left[\frac{\partial z}{\partial x}\right]^2 + \left[\frac{\partial z}{\partial y}\right]^2}\right). \quad (3)$$

The slope was calculated from the SRTM3 data using the r.slope [25] algorithm within the GRASS GIS software (version 7.4.0, Free and Open Source Software). The result is a new raster data in which each pixel is associated with the slope angle, expressed in degrees (e.g., pixels with angle values equal to 0 degrees represent a flat surface).

Another product derived from the DTM is the aspect map describing the orientation—with respect to the North—of the direction of maximum slope of the plane tangent to the ground surface.

This therefore represents the geographical exposure of the surface slope and it is obtained as the ratio, in a grid mesh, between the two main gradients of the grid along the *x*- and *y*-axis. As for the slope calculation, a 3 × 3 mobile window is applied to each cell (Figure 2) of the input raster and for each cell in the center of the window the aspect is calculated using an algorithm taking into account the values of the eight adjacent cells [26].

Consequently, the corresponding aspect is calculated with the following formula:

$$\text{aspect} = \frac{180}{\pi} \times \arctan\left(\left[\frac{\partial z}{\partial y}\right] / \left[\frac{\partial z}{\partial x}\right]\right). \quad (4)$$

The aspect is a spatial information indicating the direction or azimuth of a surface, and is measured in degrees in a range of values ranging from 0° to 360°. Values with 0°, 90°, 180°, and 270° indicate respectively East, North, West and South directions. The algorithm used for the aspect calculation is r.aspect of GRASS GIS software [25], which is based on the finite difference Horn method [24] using a 3 × 3 kernel.

2.2. Methodology

This section briefly presents the downscaling technique used to improve the spatial resolution of SRI data from AMSU-B/MHS spatial resolution to MSG-SEVIRI spatial resolution, i.e., the method of kriging with external drift, focusing on the construction of the empirical semivariogram and its fit. The trend used as auxiliary data to improve the downscaling process has been chosen according to the type of precipitation, classified into convective and stratiform.

2.2.1. Downscaling Technique

The ordinary kriging method is a downscaling technique to estimate a variable of interest, based on the spatial autocorrelation of data [27]. The value of a spatial process in a not observed site, i.e., $\hat{Z}(s_0)$ with $s_0 \notin \mathbf{s} = \{s_1, \ldots, s_n\}$, is defined as follows:

$$\hat{Z}(s_0) = \boldsymbol{\lambda} Z(\mathbf{s}), \quad (5)$$

where $\boldsymbol{\lambda}$ is a vector of real weights. The unbiasedness of $\hat{Z}(s_0)$ is satisfied by its own definition and set the sum to weights to one, i.e.,

$$\sum_{i=1}^{n} \lambda_i = 1. \quad (6)$$

The method of kriging with external drift, instead, also makes use of auxiliary information. In particular, the value of $\hat{Z}(s_0)$ is defined as follows:

$$\hat{Z}(s_0) = \mu(\mathbf{s}) + \boldsymbol{\lambda} Y(\mathbf{s}), \quad (7)$$

where $\mu(\mathbf{s})$ is called drift or trend and it is a linear combination of deterministic functions, i.e., $\mu(\mathbf{s}) = \mathbf{a} f(\mathbf{s})$ with $\mathbf{a}$ vector of coefficients, $Y(\mathbf{s}) = (Z(\mathbf{s}) - \mu(\mathbf{s}))$ represents the residual of the random process (assumed to be intrinsically stationary with mean equal to zero) and $\boldsymbol{\lambda}$ is a vector of the weights of the corresponding residual $Y(\mathbf{s})$ [27].

In both cases, the weights must be determined to minimize the estimation of the variance and to ensure the estimator has no bias. The variance estimation of the differences between the value of the variable under study at two different sites s_i and s_j [27,28] is defined through a function called semivariogram, i.e.,:

$$\gamma(s_i, s_j) = \frac{Var(Z(s_i) - Z(s_j))}{2} = \frac{E\left(|Z(s_i) - Z(s_j)|^2\right)}{2}. \quad (8)$$

The semivariogram is a function of data variability between pairs of points at various distances and describes the probabilistic properties of a process; it is very useful in the analysis of the spatial dependence of a random process. The semivariogram is a continuous function in the origin, but, in practice, it is often observed that $\gamma(0) \neq 0$ This situation is known as nugget effect ($\gamma(0) = \tau^2$ is called nugget) and it is linked to measurement errors or spatial resolution problems in sampling, which in turn affect the variogram value at very small distances. Another important feature of the semivariogram function is the sill, defined as:

$$\tau^2 + \sigma^2 = \lim_{\|h \to \infty\|} \gamma(h), \tag{9}$$

where σ^2 is known as the partial sill (the sill minus the nugget) and h is the distance between two different observations. In case the sill takes on a finite value, it means that the stochastic process is weakly stationary; in addition, if this occurs for a finite value of $h = h^*$, then h^* is said range of the variogram. The range quantifies the distance over which two different observations can be considered correlated. After the analysis of the empirical semivariogram, a fitting model must be considered to extrapolate the spatial behaviour of the observed points to the area of interest.

In the literature, there are several theoretical semivariogram models $\hat{\gamma}_i$ (e.g., linear, exponential, Gaussian, wave and circular), with known analytical properties and physical meaning of parameters. Once the empirical semivariogram function is defined and the different theoretical semivariogram models are explored, it is necessary to choose a fit criterion (Appendix A), besides the graphic inspection. These steps allow for providing the weights for the spatial interpolation through kriging techniques.

2.2.2. Choice of the Auxiliary Variables for Kriging with External Drift

A crucial step in the method of kriging with external drift lies in the selection of the auxiliary variables, which strongly depends on the parameter to be downscaled. In this work, we exploit the known relationship between rainfall and orography [1], in order to identify the most important auxiliary variables in the construction of the semivariogram. In particular, we characterize the above relationship by testing the slope, aspect and elevation data as auxiliary variables. This choice is mainly justified by the characteristics of the area under study, featuring a complex orography—i.e., a continuous alternation of mountainous areas, valleys and plains—which, in turn, affects the local climate. As precipitation may generally be classified in convective type (characterized by strong vertical velocity field, high rainfall intensity and small coverage area) and stratiform type (characterized by weak vertical velocity field, lower rainfall intensity and a more homogeneous coverage area) [29,30], we also investigated the correlation between the precipitation type and the trend to use in the method of kriging with external drift. To this aim, we took into consideration the information provided by MSG-SEVIRI, i.e., BT at four channels, in order to distinguish the cloud cover type within the AMSU/B FOV. In particular, we used:

(i) BT differences between the 10.8 μm (thermal-infrared) and the 6.2 μm (water vapour) channels, to identify deep-convection areas [11,14–17];
(ii) BT differences between the 7.3 μm (water vapour) and the 6.2 μm channels, to recognize areas of intense precipitation [17,18];
(iii) BT difference between 8.7 μm and 10.8 μm (thermal-infrared) channels, useful to discriminate liquid and ice cloud that could be associated with mid-level stratiform cloud (nimbostratus) and convective clouds, respectively.

Importantly, we have firstly performed tests using one covariate at a time, and subsequently we have used combinations of different covariates as external drift to find the combination yielding the best estimate of surface rain intensity data. The validation results for the downscaled data indicate

that two different trends should be used for the two analyzed rain types. In particular, the trend combinations for the convective (TREND_CR) and stratiform (TREND_SR) rain types are:

- TREND_CR = difference between 8.7 μm and 10.8 μm channels + slope + difference between 10.8 μm and 7.3 μm channels;
- TREND_SR = difference between 8.7 μm and 10.8 μm channels + slope + difference between 7.3 μm and 6.2 μm channels.

To notice that TREND_CR and TREND_SR are the function $\mu(\mathbf{s})$ described in Formula (7). The main difference between the two trend combinations lies in the last covariate used to single out the two types of rain. The simultaneous occurrence of convective and stratiform rain in the area under study is analysed by applying both trends on the basis of the cloud mask outcome (Classification Mask Coupling of Statistical and Physics Methods, C-MACSP [14,30]), which acts as the initial selecting criterion. In detail, when C-MACSP calls for convective type rain, the trend to be used is TREND_CR; otherwise, the trend to be used is TREND_SR. We should emphasize that the combined use of orographic features (slope, aspect and elevation) and MSG-SEVIRI data, together with the separation depending on rain type, represents a novelty element within the realm of downscaling applications by means of the kriging method with external drift. In addition, the results further support this approach, since we actually find that different covariate combinations should be used to best match the rain gauge observations.

2.2.3. Validation

The downscaled SRI data were validated against rain gauge measurements. Furthermore, to show the improvement provided by the downscaling procedures compared to the initial data, the OPEMW SRI were also compared against rain gauge observations. The correlation coefficient (*corr*), the root mean square error (*RMSE*), the mean bias error (*MBE*) and the mean absolute error (*MAE*) are used for the validation of the results obtained by kriging with external drift, considering its different drifts. These metrics are defined as follows:

$$corr = \frac{cov(p,s)}{\delta_p \delta_s}, \tag{10}$$

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(s_i - p_i)^2\right]^{1/2}, \tag{11}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(s_i - p_i), \tag{12}$$

and

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|s_i - p_i|, \tag{13}$$

where p denotes the rain gauges observations, s denotes the downscaled products, n is the number of observations, cov indicates the covariance operator, δ_p is the standard deviation of p and δ_s is the standard deviation of s.

The validation required the space-time colocation of data from the different sources. In particular, each rain gauge rain rate has been compared with the spatially closest OPEMW value and, after the downscaling procedure, with the closest downscaled value. With regard to the temporal colocation, each OPEMW SRI value is associated with the time of the satellite overpass, which is an instantaneous observation. In order to compare the rain rate measured by OPEMW SRI with rain gauge observations, we used the 15-minute cumulative rain value, as this is the shortest time interval available for all the considered stations, and convert it as to match the rain rate units (mm h^{-1}).

3. Results

The results of the validation are presented for the ten case studies listed in Table 1 These cases correspond overall to 440 comparisons between observed and both original and downscaled rain data.

Table 1. List of case studies.

Case	Data	Satellite Overpass (UTC)	Case	Data	Satellite Overpass (UTC)
1	30 Nov 2017	06:33	6	10 Feb 2018	07:54
2	28 Dec 2017	08:24	7	13 Feb 2018	06:58
3	02 Jan 2018	08:49	8	14 Feb 2018	12:13
4	10 Jan 2018	06:54	9	20 Feb 2018	18:32
5	12 Jan 2018	06:54	10	20 Mar 2018	17:43

The continuous statistical assessment, by means of *corr*, *RMSE*, *MBE* and *MAE* values, is reported in Tables 2 and 3 for each case and for the whole dataset, respectively. In detail, the continuous statistics show the comparison between rain gauge observations against the original OPEMW SRI data, the downscaled SRI data by ordinary kriging (OK) and the downscaled SRI data by kriging with external drift (KED), obtained using the trend producing the best results for each case.

Table 2. Continuous statistics against rain gauge observations for original OPEMW data, and data downscaled with ordinary kriging (OK) and kriging with external drift (KED) methods.

Case	*MAE* [mm/h]	*MBE* [mm/h]	*RMSE* [mm/h]	*corr*
	(OPEMW, OK, KED)	*(OPEMW, OK, KED)*	*(OPEMW, OK, KED)*	*(OPEMW, OK, KED)*
1	(1.46, 1.39, 1.33)	(−0.72, −1.03, −0.89)	(3.08, 2.97, 2.82)	(0.62, 0.73, 0.75)
2	(0.32, 0.35, 0.34)	(−0.22, −0.20, −0.18)	(0.69, 0.67, 0.64)	(0.24, 0.32, 0.42)
3	(1.38, 1.25, 1.03)	(0.45, 0.47, 0.41)	(2.12, 1.82, 1.46)	(0.50, 0.55, 0.72)
4	(1.17, 0.70, 0.65)	(0.42, 0.60, 0.53)	(2.73, 1.22, 1.15)	(0.62, 0.72, 0.73)
5	(2.15, 1.55, 1.43)	(0.59, 0.44, 0.38)	(3.56, 2.35, 2.48)	(0.48, 0.63, 0.68)
6	(3.29, 3.09, 2.35)	(3.29, 3.09, 2.35)	(3.92, 3.63, 2.76)	(0.33, 0.33, 0.49)
7	(1.32, 0.98, 0.92)	(0.29, 0.07, 0.04)	(1.61, 1.13, 1.06)	(0.33, 0.39, 0.41)
8	(5.21, 4.99, 4.86)	(5.21, 4.99, 4.86)	(6.12, 5.75, 5.53)	(0.27, 0.31, 0.33)
9	(1.52, 1.52, 1.52)	(−1.21, −1.39, −1.39)	(2.09, 2.03, 2.01)	(0.30, 0.31, 0.34)
10	(3.86, 3.59, 3.52)	(3.30, 3.00, 2.94)	(4.61, 4.07, 3.98)	(0.31, 0.34, 0.36)

Table 3. Continuous statistics against rain gauge observations for original OPEMW data, and data downscaled with ordinary kriging (OK) and kriging with external drift (KED) methods.

Statistics	OPEMW	OK	KED
MAE (mm/h)	2.17	1.94	1.70
MBE (mm/h)	1.14	1.44	0.91
RMSE (mm/h)	3.05	2.58	2.38
corr	0.34	0.46	0.52

The performances of the downscaling methods should be evaluated by considering also the original OPEMW SRI efficiency for the cases considered. The results show a reasonable improvement of the correlation coefficient in the comparison between rain gauge observations and downscaled data, compared to the original OPEMW SRI data. The agreement in terms of *RMSE*, *MBE* and *MAE* is better for downscaled data obtained by either of the two methods. In particular, the continuous statistics shows that kriging with external drift performs better than the ordinary kriging. In fact, the use of auxiliary information improves the estimate of the downscaled variable by +13% (*corr*), −37% (*MBE*), −8% (*RMSE*), and −12% (*MAE*). The computational cost to obtain estimates of rainfall data on the

MSG-SEVIRI grid, which in our case consists of a number of pixels of 698, is about 1.2 s. Time refers to a Intel(R) Core(TM) i5-4460 CPU 3.20GHz, with 8 Gb RAM.

Among the case studies listed in Table 1, we show in detail two representative cases (5 and 7). In Figure 3, the MSG-SEVIRI images at VIS (0.6 µm) and IR (10.8 µm) channels and the correspondent cloud classification (C-MACSP) maps are reported for both cases. The C-MACSP maps show convection in some areas of Basilicata for case 5, while C-MACSP detects optically thick clouds in the same area for case 7. Therefore, no convection is present in the second case; however, visual inspection of the temporal sequence of MSG-SEVIRI images seem to indicate midlevel clouds—under high thick clouds detected by C-MACSP—which probably generate stratiform precipitation. The visual inspection is necessary to identify multi-layered clouds because C-MACSP does not characterize multi-layered clouds yet.

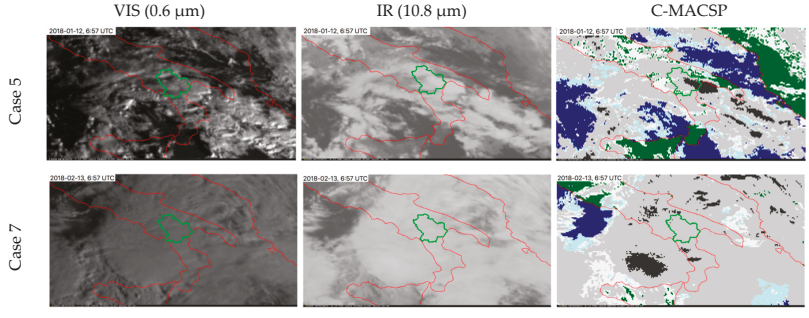


Figure 3. MSG-SEVIRI images at VIS (Visible) 0.6 µm (left column) and IR (InfraRed) 10.8 µm (middle column) and C-MACSP maps (right column) for case 5 (top row) and 7 (bottom row). C-MACSP: clear sea (blue), clear land (green), high thin cloud (cyan), low/middle cloud (white), high thick cloud (light grey) and convective cloud (dark grey).

Furthermore, Figures 4 and 5 show the comparison between OPEMW SRI data, rain gauge observations and downscaled data by kriging with external drift. The continuous statistic for the two cases considered is reported in Tables 4 and 5.

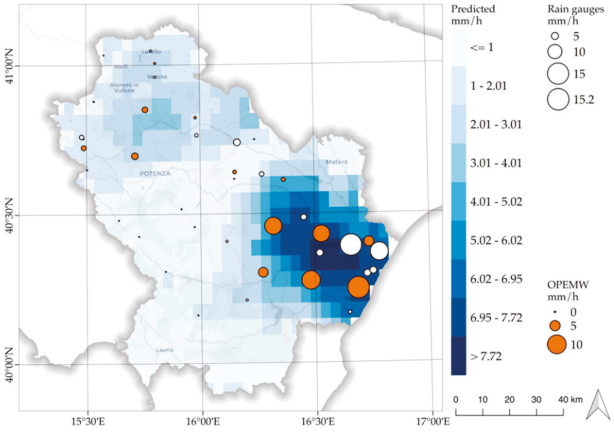


Figure 4. Comparisons between rain gauge observations, the original OPEMW SRI and downscaled data for the date 12 Jan 2018 at 06:54 a.m. UTC. The white circles represent the SRI measured by rain gauges and the orange circles represent the OPEMW SRI. The symbol size represents the SRI in mm/h.

Table 4. As in Table 3 but for data from date 12 Jan 2018 at 06:54 UTC.

Statistics	OPEMW	OK	KED
MAE (mm/h)	2.15	1.55	1.43
RMSE (mm/h)	3.56	2.48	2.35
MBE (mm/h)	0.59	0.44	0.38
corr	0.48	0.63	0.68

Considering case 5 (Figure 4), the initial OPEMW SRI data set is composed by 33 pixels; out of these, only 13 feature surface rain intensity different from zero. For this case, the results show a good improvement of correlation in the comparison between rain gauge observations and downscaled data compared to the original OPEMW SRI data (0.68 against 0.48). In addition, the agreement for downscaled data in terms of *RMSE* and *MAE* is substantial (2.35 mm/h against 3.56 mm/h and 1.43 mm/h against 2.15 mm/h, respectively).

Considering case 7 (Figure 5), the initial OPEMW SRI data set consists of 34 pixels; however, only 21 of these report surface rain intensity different from zero. In this case, the results show an improvement of correlation in the comparison between rain gauge observations and downscaled data compared to the original OPEMW SRI data (0.41 against 0.33). In addition, the agreement in terms of *RMSE* and *MAE* is significant for downscaled data (1.06 mm/h against 1.61 mm/h and 0.92 mm/h against 1.32 mm/h, respectively).

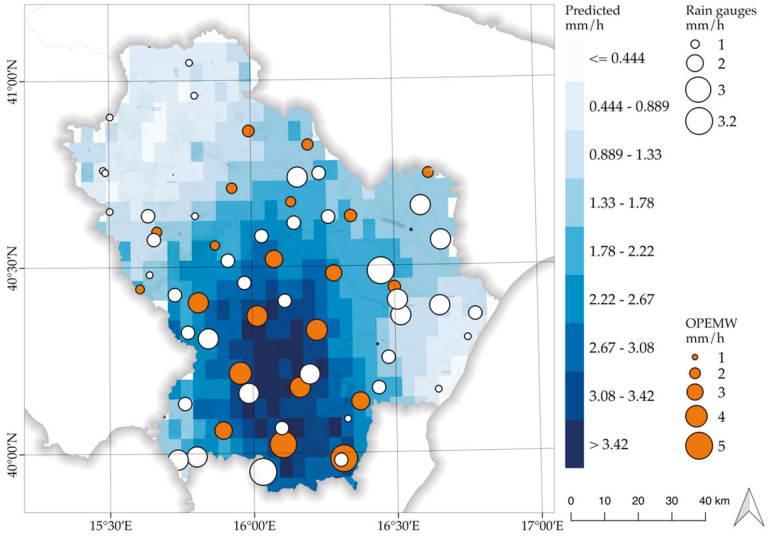


Figure 5. Comparisons between rain gauge observations, the original OPEMW SRI and the downscaled data for the date 13 Feb 2018 06:58 UTC. The white circles represent the SRI measured by rain gauges and the orange circles represent the OPEMW SRI. The symbol size represents the SRI in mm/h.

Table 5. As in Table 3 but for data from 13 Feb 2018 at 06:58 UTC.

Statistics	OPEMW	OK	KED
MAE (mm/h)	1.32	0.98	0.92
RMSE (mm/h)	1.61	1.13	1.06
MBE (mm/h)	0.29	0.07	0.04
corr	0.33	0.39	0.41

In both cases, the improvement obtained by using the auxiliary variables is evident.

4. Conclusions

In this paper, we have applied a geostatistical downscaling procedure, i.e., the method of kriging with external drift, in order to improve the spatial resolution of satellite-based rainfall observations from the original resolution to finer resolution. The kriging with external drift method features several additional advantages compared to standard kriging techniques, since it relies on auxiliary variables-sampled frequently and regularly [31]—which may further minimize the error of the estimation. Therefore, a crucial step in this method lies in the selection of the auxiliary variables, which strongly depends on the parameter to be downscaled.

The known relationship between rainfall and orography [1] lead us firstly to test the slope, aspect and elevation data as auxiliary variables. Furthermore, as precipitation may generally be classified in convective and stratiform types, we also took into account the information provided by MSG-SEVIRI BT observations, acquired in the water vapor band (6.2 μm and 7.3 μm) and in thermal-IR (10.8 μm and 8.7 μm), in order to distinguish the cloud cover and precipitation types. We initially used one covariate at a time and then a combination of different covariates as external drift to find the most suitable combination yielding the best estimate for the data analysed. In detail, kriging with external drift was applied to all the proposed cases study by varying the auxiliary variables and, subsequently, analysing the validation results for each rain type separately. From this analysis, it resulted that different auxiliary variables should be used for the two rain types. In particular, we found that the trend for the convective (TREND_CR) rain cases is a combination of slope, the 8.7 μm IR channel and the difference between 10.8 μm and 7.3 μm channels, whereas, for the stratiform (TREND_SR) rain cases, the trend is a combination of the slope, the 8.7 μm IR channel and the difference between 7.3 μm and 6.2 μm channels.

To evaluate the performances of our procedure, we considered ten case studies corresponding overall to 440 comparisons between rain gauge observations and original/downscaled rain data. The statistical analysis is based on the calculation of the *RMSE*, *MBE*, *MAE* and correlation. The results show a reasonable improvement of the correlation coefficient in the comparison between rain gauge observations and downscaled data, compared to the original OPEMW SRI data. In addition, the agreement in terms of *RMSE*, *MBE* and *MAE* is better for downscaled data. In particular, the results show that kriging with external drift clearly outperforms the ordinary kriging, which only considers coordinates and distances between observations to be downscaled. In fact, the use of auxiliary information improves the estimate of the downscaled variable by +13% (*corr*), −37% (*MBE*), −8% (*RMSE*), and −12% (*MAE*). Therefore, the proposed methodology has produced results improving the statistics compared to the original OPEMW SRI data as well as the ordinary kriging. Thus, the combination of the orographic features (slope, aspect and elevation), together with MSG-SEVIRI data, represents a novelty element that allowed us to use different covariate combination for improving the quality and resolution of satellite rainfall observations.

Author Contributions: A.C., F.R., S.L. and E.R. designed the research, wrote the paper and contributed to evaluation process. D.C., F.D.P., D.G., S.G., E.G., S.T.N., E.R. and M.V. contributed to data processing and analysis. All the co-authors helped to revise the manuscript.

Funding: This work has been financed by the Italian Ministry of Economic Development (MISE) in the framework of the SolarCloud project, contract No. B01/0771/04/X24.

Acknowledgments: The authors acknowledge the Functional Center Decentralized (FCD) of Basilicata for providing data from the rain gauge network.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A

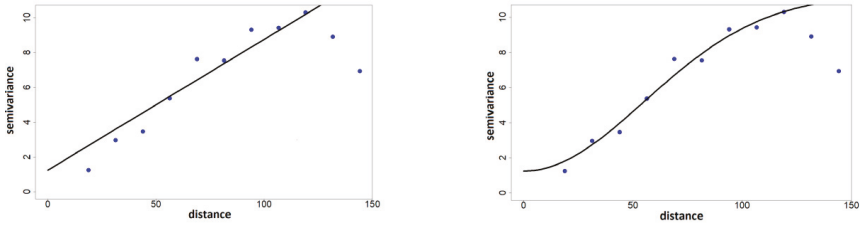
This appendix describes the analysed criteria for the choice of the theoretical semivariogram model. Given a stochastic process, several theoretical semivariogram models exist describing the features of the process, with known analytical properties and physical meaning of the parameters [31]. In the following, a theoretical semivariogram model will be denoted with $\hat{\gamma}(\mathbf{h}, \theta)$, where $\theta = \{\sigma^2, a, \tau^2\}$ is the set of the specific sill, range and nugget parameters, respectively. Once the empirical semivariogram function γ has been defined, it is necessary to choose a fit criterion for γ. In particular, the problem of finding the best $\hat{\gamma}$ for a given γ can be reduced to estimate the best set θ. In the literature, different ways are proposed to fit the empirical semivariogram to a semivariogram model, e.g., the generalized least squares methods.

In the R software, the package geoR provides xvalid, which is a function to perform model validation by comparing observed and predicted values. In particular, this function helps to determine the theoretical semivariogram that best fits the empirical semivariogram. In this work, we have also implemented a procedure for the fitting of empirical semivariograms, for comparison with the geoR tool. This procedure consists of six selection criteria that have been previously defined and tested in the context of a more complex study that involves the downscaling of different meteorological variables.

The six selection criteria choose $\hat{\gamma}$ by searching for the value of $\hat{\gamma}_i$ that minimises

1. the difference between the nugget values of γ and $\hat{\gamma}_i$;
2. the difference between the partial sills values of γ and $\hat{\gamma}_i$;
3. the difference between the range values of γ and $\hat{\gamma}_i$;
4. the RMSE between γ and $\hat{\gamma}_i$;
5. the value of the sum of the differences determined at points 1-2;
6. the value of the sum of the differences determined at points 1-2-3-4.

The statistical assessment of the six selection criteria proved that criterion 6 is the most effective and, consequently, we used it in this study. The results obtained by applying criterion 6 were compared with those obtained by using xvalid, and they generally agree except for some cases in which criterion 6 performs better as shown in the example of Figure A1.



(**a**) (**b**)

Figure A1. (**a**) the empirical semivariogram (points) and the linear semivariogram model (lines) chosen by using the xvalid function; (**b**) the empirical semivariogram (points) and the Gaussian semivariogram model (lines) chosen by using criterion 6.

Appendix B

This appendix shows some details about the considered ground observation stations.

Table A1. Identity code (ID), location and altitude of ground observation stations.

ID	Rain Gauge	Longitude (deg north)	Latitude (deg east)	Elevation (m)
1	ABRIOLA A SELLATA PIERFAONE	15.76106	40.50064	1463
2	ALBANO DI LUCANIA	16.03541	40.58202	758
3	ANZI SIMN	15.91608	40.5167	929
4	AVIGLIANO	15.47889	40.75972	592
5	BALVANO	15.5015	40.64956	385
6	BASENTO FREATIMETRO	16.78119	40.36838	9
7	BRADANO PONTE COLONNA	16.16268	40.73881	213
8	BRIENZA	15.64131	40.47969	786
9	CASTELSARACENO PC	15.98547	40.16081	1090
10	CASTROCUCCO	15.80183	39.99217	142
11	CAVONE SS106	16.72739	40.29586	13
12	CRACO PESCHIERA	16.52011	40.36642	57
13	EPISCOPIA	16.09883	40.0667	578
14	FERRANDINA SP	16.45156	40.48611	457
15	GORGOGLIONE	16.11419	40.40786	1051
16	GRASSANO SP	16.27011	40.63236	486
17	GRUMENTO-PONTE LA MARMORA	15.84508	40.30835	552
18	IRSINA PC	16.23947	40.74858	552
19	LAGONEGRO PC	15.76206	40.13419	791
20	LAURENZANA	15.97322	40.45678	814
21	LAVELLO	15.78608	41.04806	304
22	MARATEA MASSA	15.73597	39.98358	492
23	MARSICO NUOVO PC	15.72939	40.4265	747
24	MATERA	16.59539	40.65969	403
25	MONTESCAGLIOSO SIMN	16.66371	40.56673	162
26	MURO LUCANO	15.48673	40.75361	586
27	NOEPOLI	16.32989	40.08975	556
28	OFANTO A MONTICCHIO	15.50351	40.90276	322
29	OPPIDO LUCANO	15.98543	40.76388	747
30	PICERNO	15.63724	40.63771	655
31	POTENZA	15.80161	40.63703	820
32	ROCCANOVA	16.19922	40.21056	704
33	ROTONDA SIMN	16.03237	39.95003	557
34	SINNI A VALSINNI	16.4399	40.17283	152
35	SINNI SS106	16.64803	40.16556	15
36	STIGLIANO	16.50908	40.40836	150
37	TERRA MONTONATA	16.75283	40.30469	7
38	TERRANOVA DI POLLINO SIMN	16.30374	39.97981	936
39	TITO	15.65703	40.57425	661
40	TORRE ACCIO	16.65694	40.39072	19
41	TRAMUTOLA	15.77394	40.32528	662
42	TRICARICO SIMN	16.14868	40.61671	682
43	TURSI	16.47469	40.25375	264
44	VENOSA	15.80325	40.95986	430

References

1. Cantet, P. Mapping the mean monthly precipitation of a small island using kriging with external drifts. *Theor. Appl. Climatol.* **2017**, *127*, 31–44. [CrossRef]
2. Levizzani, V.; Bauer, P.; Turk, F.J. Measuring precipitation from space. In *EURAINSAT and the Future*; Springer: Dordrecht, The Netherlands, 2007; 748p, ISBN 978-1-4020-5834-9.
3. Guan, H.; Wilson, J.L.; Xie, H.J. A cluster-optimizing regression-based approach for precipitation spatial downscaling in mountainous terrain. *J. Hydrol.* **2009**, *375*, 578–588. [CrossRef]
4. 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]
5. Immerzeel, W.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]

6. Jia, S.; Zhu, W.; Lu, A.; Yan, T. A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in theQaidam Basin of China. *Remote Sens. Environ.* **2011**, *115*, 3069–3079. [CrossRef]
7. 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]
8. Alexakis, D.D.; Tsanis, I.K. 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]
9. Di Tomaso, E.; Romano, F.; Cuomo, V. Rainfall estimation from satellite passive microwave observations in the range 89 GHz to 190 GHz. *J. Geophys. Res.* **2009**, *114*, D18203. [CrossRef]
10. Cimini, D.; Romano, F.; Ricciardelli, E.; Di Paola, F.; Viggiano, M.; Marzano, F.S.; Colaiuda, V.; Picciotti, E.; Vulpiani, G.; Cuomo, V. Validation of satellite OPEMW precipitation product with ground-based weather radar and rain gauge networks. *Atmos. Meas. Tech.* **2013**, *6*, 3181–3196. [CrossRef]
11. Schmetz, J.; Pili, P.; Tjemkes, S.; Just, D.; Kerkmann, J.; Rota, S.; Ratier, A. An Introduction to Meteosat Second Generation (MSG). *Bull. Am. Meteorol. Soc.* **2002**, *83*, 977–992. [CrossRef]
12. Nilo, S.T.; Romano, F.; Cermak, J.; Cimini, D.; Ricciardelli, E.; Cersosimo, A.; Di Paola, F.; Gallucci, D.; Gentile, S.; Geraldi, E.; et al. Fog Detection Based on Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager High Resolution Visible Channel. *Remote Sens.* **2018**, *10*, 541. [CrossRef]
13. Gallucci, D.; Romano, F.; Cersosimo, A.; Cimini, D.; Di Paola, F.; Gentile, S.; Geraldi, E.; Larosa, S.; Nilo, S.T.; Ricciardelli, E.; et al. Nowcasting Surface Solar Irradiance with AMESIS via Motion Vector Fields of MSG-SEVIRI Data. *Remote Sens.* **2018**, *10*, 845. [CrossRef]
14. Ricciardelli, E.; Cimini, D.; Di Paola, F.; Romano, F.; Viggiano, M. A statistical approach for rain intensity differentiation using Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager observations. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2559–2576. [CrossRef]
15. 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]
16. Ricciardelli, E.; Di Paola, F.; Gentile, S.; Cersosimo, A.; Cimini, D.; Gallucci, D.; Geraldi, E.; Larosa, S.; Nilo, S.T.; Ripepi, E.; et al. Analysis of Livorno Heavy Rainfall Event: Examples of Satellite-Based Observation Techniques in Support of Numerical Weather Prediction. *Remote Sens.* **2018**, *10*, 1549. [CrossRef]
17. Di Paola, F.; Ricciardelli, E.; Cimini, D.; Romano, F.; Viggiano, M.; Cuomo, V. Analysis of Catania Flash Flood Case Study by Using Combined Microwave and Infrared Technique. *J. Hydrometeorol.* **2014**, *1*. [CrossRef]
18. Barbosa, H.A.; Ertürk, A.G.; Da Silva, L.R.M. Using the Meteosat-9 images to the detection of deep convective systems in Brazil. *J. Hyperspectr. Remote Sens.* **2012**, *12*, 069–082. [CrossRef]
19. Berterame, D.; D'Avenia, G.; Glisci, C.; Lanorte, V.; Mangiolfi, A.; Motta, G.; Pacifico, G. Analisi del Deficit Pluviometrico del 1° Semestre 2017 in Basilicata. 2017. Available online: http://www.centrofunzionalebasilicata.it/ew/ew_pdf/r/Articolo_Deficit_I_semestre_2017_colonne.pdf (accessed on 6 November 2018).
20. Farr, T.G.; Kobrick, M. Shuttle Radar Topography Mission produces a wealth of data. *Eos. Trans. Am. Geophys. Union* **2000**, *81*, 583–585. [CrossRef]
21. Rabus, B.; Eineder, M.; Roth, A.; Bamler, R. The shuttle radar topography mission—A new class of digital elevation models acquired by spaceborne radar. *Remote Sens.* **2003**, *57*, 241–262. [CrossRef]
22. Werner, M. Shuttle radar topography mission (SRTM), mission overview. *J. Telecommun. (Frequenz)* **2001**, *55*, 75–79. [CrossRef]
23. SRTM 90 m Digital Elevation Database v4.1. Available online: ftp://e0srp01u.ecs.nasa.gov (accessed on 6 November 2018).
24. Horn, B.K.P. Hill Shading and the Reflectance Map. *Proc. IEEE* **1981**, *69*, 14–47. [CrossRef]
25. Hofierka, J.; Mitasova, H.; Neteler, M. Geomorphometry in GRASS GIS. In *Geomorphometry: Concepts, Software, Applications. Developments in Soil Science*; Hengl, T., Reuter, H.I., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 33, pp. 387–410.
26. Burrough, P.; McDonnell, R.A. *Principles of Geographical Information Systems*; Oxford University Press: Oxford, UK, 1998.
27. Cressie, N.A.C. Statistics for spatial data, revised edition. *Wiley Ser. Probab. Stat.* **2015**. [CrossRef]
28. Wackernagel, H. *Multivariate Geostatistics. An Introduction with Applications*; Springer: Berlin, Germany, 2003.

29. 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]
30. Ricciardelli, E.; Romano, F.; Cuomo, V. Physical and statistical approaches for cloud identification using Meteosat Second Generation-Spinning Enhanced Visible and Infrared Imager Data. *Remote Sens. Environ.* **2008**, *112*, 2741–2760. [CrossRef]
31. Karl, J.K. Spatial predictions of cover attributes of rangeland ecosystems using regression kriging and remote sensing. *Rangeland Ecol. Manag.* **2010**, *63*, 335–349. [CrossRef]

remote sensing

MDPI

Article

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

Cheng Chen [2], Qiuwen Chen [1,2,*], Zheng Duan [3], Jianyun Zhang [1], Kangle Mo [2], Zhe Li [4] and Guoqiang Tang [5]

1 State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing 210029, China; jyzhang@nhri.cn
2 Center for Eco-Environmental Research, Nanjing Hydraulic Research Institute, Nanjing 210029, China; chencheng@nhri.cn (C.C.); klmo@nhri.cn (K.M.)
3 Chair of Hydrology and River Basin Management, Technical University of Munich, Arcisstrasse 21, 80333 Munich, Germany; duanzheng2008@gmail.com
4 Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; lizhe@igsnrr.ac.cn
5 State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China; 15201514761@163.com
* Correspondence: qwchen@nhri.cn; Tel.: +86-25-85829765

Received: 16 May 2018; Accepted: 11 June 2018; Published: 14 June 2018

Abstract: The performance of the latest released Integrated Multi-satellitE Retrievals for GPM mission (IMERG) version 5 (IMERG v5) and the TRMM Multisatellite Precipitation Analysis 3B42 version 7 (3B42 v7) are evaluated and compared at multiple temporal scales over a semi-humid to humid climate transition area (Huaihe River basin) from 2015 to 2017. The impacts of rainfall rate, latitude and elevation on precipitation detection skills are also investigated. Results indicate that both satellite estimates showed a high Pearson correlation coefficient (*r*, above 0.89) with gauge observations, and an overestimation of precipitation at monthly and annual scales. Mean daily precipitation of IMERG v5 and 3B42 v7 display a consistent spatial pattern, and both characterize the observed precipitation distribution well, but 3B42 v7 tends to markedly overestimate precipitation over water bodies. Both satellite precipitation products overestimate rainfalls with intensity ranging from 0.5 to 25 mm/day, but tend to underestimate light (0–0.5 mm/day) and heavy (>25 mm/day) rainfalls, especially for torrential rains (above 100 mm/day). Regarding each gauge station, the IMERG v5 has larger mean *r* (0.36 for GPM, 0.33 for TRMM) and lower mean relative root mean square error (*RRMSE*, 1.73 for GPM, 1.88 for TRMM) than those of 3B42 v7. The higher probability of detection (*POD*), critical success index (*CSI*) and lower false alarm ratio (*FAR*) of IMERG v5 than those of 3B42 v7 at different rainfall rates indicates that IMERG v5 in general performs better in detecting the observed precipitations. This study provides a better understanding of the spatiotemporal distribution of accuracy of IMERG v5 and 3B42 v7 precipitation and the influencing factors, which is of great significance to hydrological applications.

Keywords: GPM IMERG v5; TRMM 3B42 v7; precipitation; evaluation; Huaihe River basin

1. Introduction

Precipitation is a major force in global climate change, hydrological cycles, and ecological environments, which is of great importance to agricultural irrigation and disaster prevention [1–3].

Precipitation data at high spatiotemporal resolutions are essential for hydrological simulation at local, regional, and global scales [4] as the key component of the global water cycle. Conventional rain gauge observations at the basin scale are usually relatively sparse and are insufficient for accurately characterizing the spatial distribution of precipitation. In contrast, satellite remote sensing has the advantages of completely scanning the entire study region and convenient access to the data, providing an alternate way to monitor precipitation at regional and global scales.

More and more global precipitation datasets have been produced with the development of various spaceborne and related satellite-based precipitation retrieval algorithms, including the Precipitation Estimation from Remote Sensing Information using Artificial Neural Network (PERSIANN) [5], Climate Prediction Center Morphing (CMORPH) [6], Global Satellite Mapping of Precipitation (GSMaP) [7], Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) [8], TRMM Multisatellite Precipitation Analysis data (TMPA) [9], and Integrated Multi-satellitE Retrievals for GPM mission (IMERG) [10].

Building upon the success of TRMM, the GPM mission is expected to advance the understanding of Earth's water and energy cycle, improve forecasting of extreme rainfall events, and extend current capabilities for using accurate and timely information on precipitation to directly benefit society [10]. Compared with TRMM, GPM provides precipitation measurements at a higher spatial (from 0.1° to 0.25°) and temporal (from half-hourly to 3-hourly) resolution. GPM offers orbital and gridded products at three different levels of data processing. The Level-3 product, the Integrated Multi-satellitE Retrievals for GPM (IMERG), combines GMI–DPR rainfall averages and rainfall estimates with data from all active and passive microwave instruments in the GPM constellation [11]. The IMERG product is available in three distinct modes—"Early Run" (latency: ~6 h after observation time), "Late Run" (latency: ~18 h) and "Final Run" (latency: ~4 months), depending upon latency and application requirements. IMERG Final Run is not only the longest daily precipitation dataset among the IMERG products, but also incorporates monthly precipitation-gauge analyses, which provides more accurate results [12]. Recently, the assessment of the IMERG Day 1 Final Run product has been carried out in various regions of the world.

The earliest evaluation and comparison of GPM Day-1 IMERG version 4 (IMERG v4) and TRMM 3B42 v7 products was carried out over Mainland China from April to December 2014 at the hourly time scale at multiple spatiotemporal scales by Tang et al. [13]. The results showed that Day-1 IMERG v4 better captures the diurnal variability of precipitation, but both products have room for further improvement, particularly in the dry climate and high-latitude regions. Tang et al. [14] also conducted a comparison study of GPM IMERG v4 and TRMM 3B42 v7 with a dense gauge network over the mid-latitude Ganjiang River basin in southeast China from May to September 2014 (rainy season) using the Coupled Routing and Excess Storage (CREST) hydrologic model. It was shown that the daily IMERG v4 product could adequately substitute for the 3B42 v7 products both statistically and hydrologically. Ma et al. [15] analyzed the three-hourly rainfall estimates over the Tibetan Plateau in the warm season of 2014. It was found that both IMERG v4 and 3B42 v7 showed similar rainfall patterns and captured the northward dynamic life cycle of the Indian monsoon reasonably well, and that IMERG v4 showed the potential for detecting solid precipitation, which cannot be retrieved from the 3B42 v7 products. Gaonao et al. [16] evaluated the first full year (12 March 2014 to 1 April 2015) performance of IMERG v4 against radar rainfall fields at 30 min, 24 h, monthly, and yearly durations above the land surface of the Netherlands, where the TRMM products were not available. Similar studies have also been carried out in India [17,18], Southern Canada [19] and Singapore [20].

GPM provides more detailed information on precipitation with higher spatial and temporal resolution than TRMM as a new generation of global precipitation observations. Since most gridded precipitation products have global (or quasi-global) coverage, their performance is expected to vary among regions [21,22]. To the authors' knowledge, studies of satellite precipitation in climatic transition zones still remain insufficient [23]. Sun et al. [24] had taken the Huaihe River basin, a typical semi-humid to humid climate transition area, as a study case to evaluate four satellite–gauge merged

quantitative precipitation estimation products (TRMM 3B42 v7, CMORPH BLD, CMORPH CRT and CMORPH CMA) during 2003–2012 and further applied them in a distributed Variable Infiltration Capacity (VIC) model to assess their hydrologic applicability. Additionally, the current study still remains on the IMERG v4. The latest IMERG Final Run product version 5 (abbreviated as IMERG v5 hereafter), which has been available since March 2014, was released in November 2017. Compared with IMERG v4, the IMERG v5 data has made several important improvements [25]. The GPROF-TMI is not presently included in the IMERG Final datasets (version 5), for example. Therefore, evaluation of these precipitation products, especially the latest released GPM IMERG v5 precipitation product with reliable measurements in climatic transition zones, is an important topic, which is beneficial to both product developers and users.

This study assesses GPM-era and TRMM-era precipitation estimates/products (IMERG v5 and 3B42 v7) over a typical semi-humid to humid climate transition zone of China. The objectives are to (1) evaluate and compare the performance of IMERG v5 and 3B42 v7 precipitation products at multiple scales in the climatic transition basin (Huaihe River basin), (2) analyze the precipitation detection capability of 3B42 v7 and IMERG v5 satellite sensors at different rainfall rate thresholds, and (3) investigate the influence of local topographical features on precipitation retrievals. The study will contribute to quantitative assessment of the potential applicability of satellite precipitation products in hydrological prediction and disaster prevention in climatic transition regions.

2. Materials and Methods

2.1. Study Area

The Huaihe River basin is located in the east of China at latitude 30°55′ N–36°36′ N and longitude 111°55′ E–121°25′ E and is about halfway between the Yellow River and Yangtze River basins (Figure 1), with a total watershed area of 270,000 km^2. It is in the middle of the North-South climatic transition zone of China, which belongs to a typical semi-humid to humid climate. It is surrounded on three sides by mountains, on the east by the sea, with vast plains in the middle. The plain area accounts for 75% of the watershed, and the mountains and hills are distributed in the west, southwest, and northeast. The Huaihe is the dividing line of the Huaihe River basin, and its northern and southern parts belong to the temperate monsoon climate area and subtropical monsoon climate, respectively. There are many tributaries and large lakes, such as Hongze Lake, Nansi Lake, Luoma Lake, and Gaoyou Lake, in the Huaihe River basin (Figure 1). Rainfall in the Huaihe River basin is relatively abundant but unevenly distributed as one of China's representative extreme rainstorm-prone areas. Generally, basin-scale average annual precipitation is about 920 mm/year, and precipitation is mainly concentrated during the rainy period from May to October. The basin has been suffering from frequent floods and droughts. Therefore, timely and accurate monitoring of the spatiotemporal distribution of regional rainfall plays an important role in hydrological and meteorological predictions and natural hazard alarms, especially in the climatic transition region.

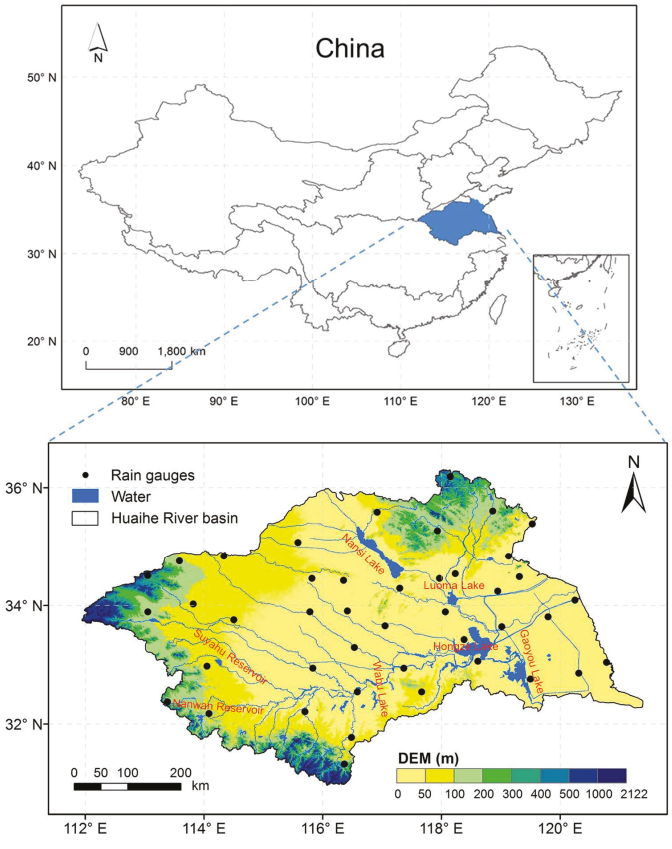


Figure 1. Geographical location, DEM and the distribution of rain gauges in the Huaihe River basin.

2.2. Datasets and Preprocessing

2.2.1. TRMM 3B42 v7

The Tropical Rainfall Measuring Mission (TRMM) launched on 27 November 1997 is a joint space mission between the National Aeronautics and Space Administration (NASA) and the Japanese Aerospace Exploration Agency (JAXA), which is designed to monitor and study tropical rainfall. The TRMM Multi-satellite Precipitation Analysis (TMPA) is one of the TRMM products designed to combine precipitation measurements from various satellite systems and precipitation gauges [9], which consist of two main products, the 3-hourly combined microwave-IR estimates (3B42) and monthly combined microwave-IR-gauge (3B43). The 3B42 precipitation estimates, which have a 3-hourly temporal resolution and a 0.25° spatial resolution with spatial coverage extending from 50° S to 50° N, have been widely used in hydrometeorological research during the past decade [11].

The version-7 post-real-time 3B42 (abbreviated as 3B42 v7) estimates from 2015 to 2017 were used, and the original 3-hourly data was accumulated to obtain daily precipitation. The 3B42 v7 data were downloaded from the Precipitation Measurement Missions (PMM) website (http://pmm.nasa.gov). Major changes were made to the algorithm to involve the radar reflectivity—rainfall rate relationship, surface clutter detection over high terrain, a new reference database for the passive microwave algorithm, and a higher-quality gauge analysis product for monthly bias correction with respect to its

predecessor, Version 6 [26,27]. The 3B42 v7 data were resampled to the same spatial resolution (0.1°) as IMERG v5 data by using the standard bilinear interpolation method [15,28] to make them comparable. The 3B42 v7 daily precipitation data were accumulated to obtain the monthly and annual precipitation.

2.2.2. GPM IMERG v5

The Global Precipitation Measurement (GPM) Core Observatory initiated by NASA and JAXA was launched on 27 February 2014. GPM, which consists of a Ku/Ka-band Dual-frequency Precipitation Radar (DPR) and a multi-channel GPM Microwave Imager (GMI), is expected to provide the next-generation global observations of rain and snow and to improve weather and precipitation forecasts through assimilation of instantaneous precipitation information. The Integrated Multi-satellitE Retrievals for GPM (IMERG) is the Level 3 multi-satellite precipitation algorithm of GPM, which combines intermittent precipitation estimates from all constellation microwave sensors, IR-based observations from geosynchronous satellites, and monthly gauge precipitation data [10,11]. There are three different daily IMERG products, which are IMERG Day 1 Early Run (near real-time with a latency of 6 h), IMERG Day 1 Late Run (reprocessed near real-time with a latency of 18 h) and IMERG Day 1 Final Run (gauged-adjusted with a latency of four months) [29]. The IMERG Final Run product provides more accurate precipitation information than the near real-time products across GPCC-gauged regions [11].

The latest released GPM IMERG version 5 (IMERG v5) dataset is produced at NASA Goddard Earth Sciences (GES) 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 IMERG precipitation products have a relatively finer spatial (0.1° latitude/longitude) and temporal (half-hourly) resolution, with spatial coverage from 60° S to 60° N. The calibrated IMERG Day 1 Final Run v5 data at daily scale over the Huaihe River basin from 2015 to 2017 were used in the current study, and the daily precipitation data were accumulated to obtain the monthly and annual precipitation. The GPM IMERG v5 data were also downloaded from the PMM website.

2.2.3. Rain Gauge Data

A dataset with daily rain gauge measurements in the Huaihe River basin from 2015 to 2017 was used as the ground reference for the evaluation of TRMM 3B42 v7 and GPM IMERG v5. The dataset contained the observations from 43 rain gauge stations of China's ground climatological data sheet (V3.0) (SURF_CLI_CHN_MUL_DAY_V3.0) and was provided by the China Meteorological Data Service Center (http://www.cma.gov.cn) [30,31]. The SURF_CLI_CHN_MUL_DAY_V3.0 dataset contains data from 824 benchmark and basic weather stations across the country, providing high quality information on precipitation since January 1951. All data were compiled by the National Meteorological Information Center and subjected to a series of quality controls, including examining extreme values and an internal consistency check, and removal of questionable data according to the methods described by Xu et al. [32]. The locations of the rain gauges are shown in Figure 1.

2.2.4. DEM Data

The Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) data (http://www.nasa.gov/) were used to analyze the impact of elevation on the spatial distribution of precipitation in this study. The original 90 m SRTM DEM data was resampled to 0.1° spatial resolution to make it consistent with the evaluated satellite precipitation data.

2.3. Methodologies

To evaluate the performance of TRMM 3B42 v7 and GPM IMERG v5 precipitation products and ensure proper error analysis on a grid basis at multiple scales over the basin, the specified location or pixel for each rain gauge was checked, and then, the spatial–temporal coincident satellite pixel with corresponding gauges was evaluated [15,20]. Only the satellite grids that covered at least one rain

gauge were evaluated, and the other grids with no rain gauges were excluded from the evaluation. There were no pixels with more than one rain gauge. The evaluation and comparison were also conducted at each individual gauge station, and the Thiessen polygon method was used to divide the basin into 43 polygons based on gauge stations. Additionally, the probability distribution function (PDF) and mean daily precipitation were calculated at different rainfall rates, and the impacts of latitude and elevation on precipitation were also assessed. It should be emphasized that the gauge observations at daily scale with no precipitation were excluded from this evaluation [15].

Several evaluation indicators were selected, including Pearson correlation coefficients (r), relative bias (RB), and relative root mean square error ($RRMSE$). The r was used to quantify the type of correlation and dependence between satellite products and gauge observations in fundamental statistics. The RB and $RRMSE$ were used to describe the bias and error of satellite precipitation compared with gauge observations [33–35]. The equations of the indicators are as follows:

$$r = \frac{\sum_{i=1}^{n}(P_i - \overline{P})(G_i - \overline{G})}{\sqrt{\sum_{i=1}^{n}(P_i - \overline{P})^2}\sqrt{\sum_{i=1}^{n}(G_i - \overline{G})^2}}, \quad (1)$$

$$RB = \frac{\sum_{i=1}^{n}(P_i - G_i)}{\sum_{i=1}^{n} G_i}, \quad (2)$$

$$RRMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - G_i)^2}{n}} / (\frac{1}{n}\sum_{i=1}^{n} G_i), \quad (3)$$

where n represents the number of rain gauge stations; P_i and G_i represent satellite precipitation estimates and rain gauge observed precipitation (mm), respectively; $\overline{P}$ and $\overline{G}$ represent mean satellite precipitation estimates and mean rain gauge observed precipitation (mm), respectively.

The probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI) were calculated at a daily time step to quantitatively examine the potential of satellite products for precipitation detection at different rainfall thresholds. The FAR gave the fraction of spurious events among all the events the satellites detected [36,37]. The POD indicated the fraction of precipitation events that were correctly detected, whereas the CSI denoted the number of correct forecasts of a rain event, divided by the total of hits, false alarms, and misses [38]. The 2 × 2 contingency table for each precipitation threshold is shown in Table 1, where hits (H) represent precipitation observed by the rain gauge and satellite simultaneously, false alarms (F) represent precipitation observed by the satellite but not observed by the rain gauge, misses (M) represent precipitation observed by the rain gauge but not observed by the satellite, and correct negatives (CN) represents precipitation observed neither by the rain gauge nor the satellite. Eight thresholds (0.5, 1, 2, 5, 10, 25, 50, 100 mm/day) were defined to identify the capability of 3B42 v7 and IMERG v5 precipitation data at different rainfall rates in this study. The POD, FAR, and CSI at each threshold were calculated by Equations (4)–(6), and the satellite precipitation data showed perfect agreement with rain gauge observations when $FAR = 0$, $POD = 1$, and $CSI = 1$.

$$POD = \frac{H}{H + M}, \quad (4)$$

$$FAR = \frac{F}{H + F}, \quad (5)$$

$$CSI = \frac{H}{H + M + F}, \quad (6)$$

Table 1. The 2 × 2 contingency table summarizing the agreement between rain gauges (G) and satellite precipitation products (P) for each precipitation threshold.

	G ≥ Threshold	G < Threshold
P ≥ threshold	H (hits)	F (false alarms)
P < threshold	M (misses)	CN (correct negatives)

3. Results

3.1. Spatiotemporal Assessment of IMERG v5 and 3B42 v7 Precipitation Estimates

Figure 2a,c shows the scatterplots of TRMM 3B42 v7 and GPM IMERG v5 satellite precipitation products versus rain gauge data at daily, monthly, and annual scales from 2015 to 2017 over the Huaihe River basin, respectively. The evaluation metrics were also included in Figure 2. Both satellite precipitation products showed low accuracy at daily scale with small *r* (0.314 for 3B42 v7, 0.413 for IMERG v5). The values of *r* significantly increased from daily scale to monthly scale (0.898 and 0.914 for 3B42 v7 and IMERG v5, respectively) and annual scale (0.896 and 0.939 for 3B42 v7 and IMERG v5, respectively). The *RRMSE* of the two satellite precipitation products was decreasing as the time scale increased. The positive values of *RB* at monthly and annual scales indicated that satellite precipitation products tended to overestimate the gauge precipitation in the Huaihe River basin, while the negative values of *RB* at daily scale indicated an underestimation. Compared with the 3B42 v7 data, the IMERG v5 data with larger *r* and lower *RRMSE* had better performance.

Table 2 shows the evaluation results (*r*, *RB* and *RRMSE*) of TRMM 0.25°, TRMM 0.1°, and GPM 0.1° rainfall estimates at multiple timescales from 2015 to 2017 over the Huaihe River basin. TRMM 0.25° means that the initial TRMM 3B42 v7 data (0.25°) was evaluated against the gauge stations, while TRMM 0.1° means the interpolated TRMM 3B42 v7 data (0.1°) was used for the evaluation. It is seen in Table 2 that at daily scale, the evaluation results of TRMM 0.25° were the worst with the smallest *r* (0.259), the largest *RRMSE* (1.965), and largest absolute *RB* (0.385). Regarding monthly and annual scales, the evaluation results of TRMM 0.25° were slightly better than that of TRMM 0.1°. Overall, GPM IMERG data (0.1°) had the best performance. It must be noted that all the statistical analyses of TRMM data afterwards were based on the 0.1° interpolated TRMM 3B42 v7 data to make it comparable with GPM IMERG v5 data.

Table 2. Evaluation results of TRMM 0.25°, TRMM 0.1°, and GPM 0.1° precipitation estimates at multiple timescales from 2015 to 2017 over the Hauihe River basin.

Time Scales	Satellite Products	*r*	*RB*	*RRMSE*
Daily	TRMM 0.25°	0.259	–0.385	1.965
	TRMM 0.1°	0.314	–0.273	1.897
	GPM 0.1°	0.413	–0.176	1.753
Monthly	TRMM 0.25°	0.904	0.059	0.442
	TRMM 0.1°	0.896	0.063	0.453
	GPM 0.1°	0.914	0.117	0.424
Annual	TRMM 0.25°	0.901	0.059	0.141
	TRMM 0.1°	0.896	0.063	0.148
	GPM 0.1°	0.939	0.117	0.156

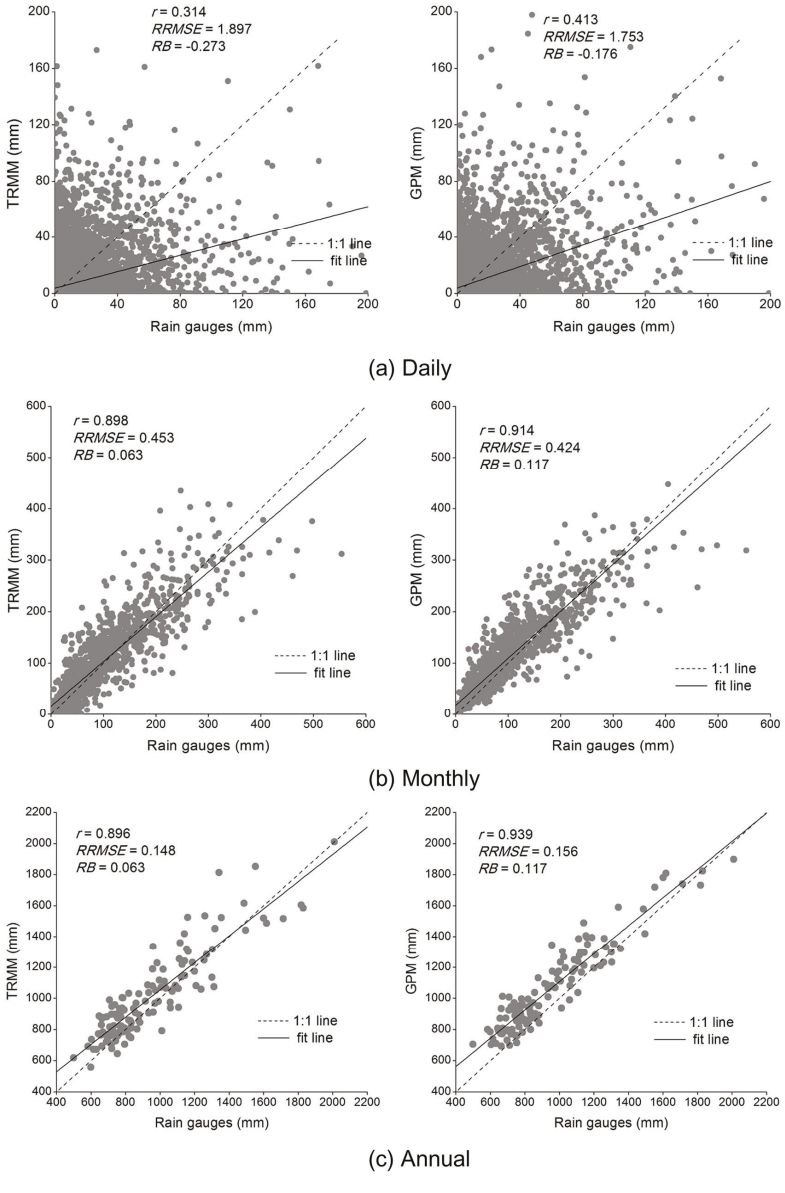


Figure 2. Scatterplots for TRMM 3B42 v7 and GPM IMERG v5 precipitation versus gauge measurements from 2015 to 2017 over the Huaihe River basin: (**a**–**c**) daily, monthly, and annual scales, respectively.

Figures 3 and 4 show the time series of *RRMSE* for mean daily and monthly precipitation from 2015 to 2017 over the Huaihe River basin, respectively. The *RRMSE* was calculated by comparing the satellite precipitation products against the corresponding 43 rain gauges data at daily and monthly scales. The mean daily and monthly precipitation values were calculated by averaging all observations from the 43 stations in the study area. It is seen in Figure 3 that the mean daily precipitation was large during summer (June, July and August) and was small during winter (December, January,

and February). The relatively large values of *RRMSE* occurred in October and December of 2016. Figure 4 demonstrates that the monthly precipitation was concentrated mainly from May to October, except the precipitation in November of 2015 and October of 2016 was abnormally large. Both satellite products agreed very well with the rain gauge data at monthly scale. The large *RRMSE* occurred mainly in winter, while the low *RRMSE* occurred in summer. Overall, the *RRMSE* of IMERG v5 was lower than that of 3B42 v7, especially in winter, which demonstrated the IMERG v5 data had better accuracy in the Huaihe River basin.

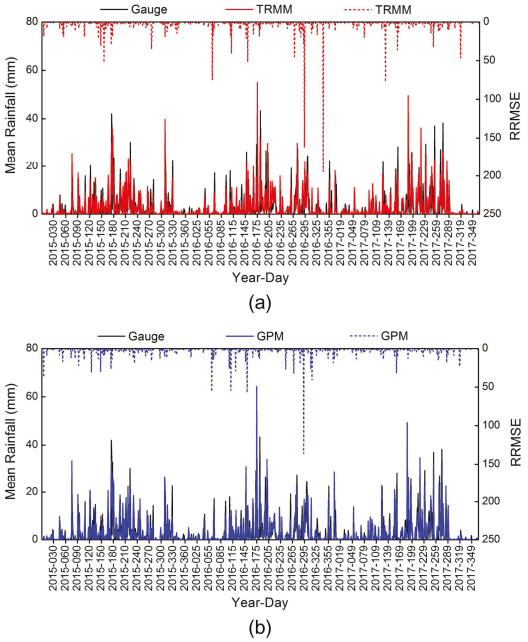


Figure 3. Time series of *RRMSE* (dash line) and mean daily precipitation (solid line) from 2015 to 2017 over the Huaihe River basin: (**a**) TRMM 3B42 v7, (**b**) GPM IMERG v5.

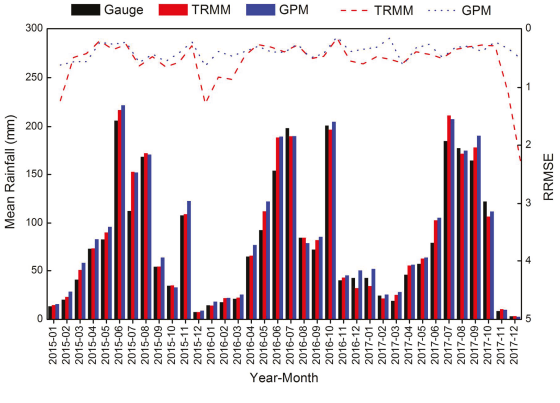


Figure 4. Time series of *RRMSE* (dot/dash line) and mean monthly precipitation (bar graph) from 2015 to 2017 over the Huaihe River basin.

Figure 5 shows the spatial distribution of mean daily precipitation from 2015 to 2017 retrieved from TRMM 3B42 v7 and GPM IMERG v5 in the Huaihe River basin. The spatial patterns of two satellite precipitation products were similar, with precipitation decreasing from the southeast to the northwest. It was found that rainfall greater than 4.0 mm/day mainly occurred in the south and southeast of the Huaihe River basin, while the relatively small rainfall occurred in the west and north. A distinct dividing line appears with the mean daily precipitation value of about 2.2 mm/day (approximately 800 mm/year), which could be explained as the geographical boundary between China's southern and northern regions. Additionally, the abnormal blue pixels in Nansi lake, Luoma lake, Hongze lake, and Nanwan reservoir in Figure 5 show that the TRMM 3B42 v7 precipitation exhibited significant overestimation over water bodies, while the GPM IMERG v5 precipitation provided more consistent precipitation estimates with much lower systematic overestimation in these areas due to the employment of a unified and updated PMW retrieval algorithm (GPROF V05).

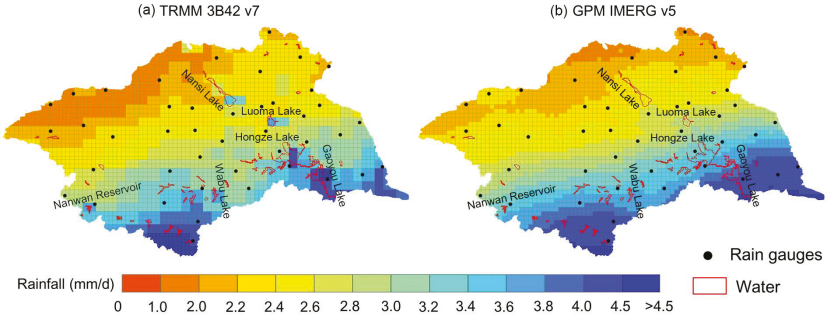


Figure 5. Spatial distribution of TRMM 3B42 v7 (**a**) and GPM IMERG v5 (**b**) mean daily precipitation from 2015 to 2017 in the Huaihe River basin.

3.2. Evaluation and Comparison at Each Individual Gauge Station

Figure 6 shows the point-based evaluation results of TRMM 3B42 v7 and GPM IMERG v5 daily gridded-precipitation retrievals against each rain gauge over the Huaihe River basin from 2015 to 2017. The *r* varied widely at each individual station and was 0.06–0.58 for 3B42 v7 and 0.14–0.64 for IMERG v5, respectively. The mean values of *r* for IMERG v5 were slightly better than that of 3B42 v7, which indicated that the GPM precipitation data had a higher consistency with gauge station records. The mean values of *RRMSE* were 1.88 for 3B42 v7 and 1.73 for IMERG v5. Both satellite precipitation products had negative *RB*, demonstrating that the underestimation of satellite precipitation occurred at all gauge stations.

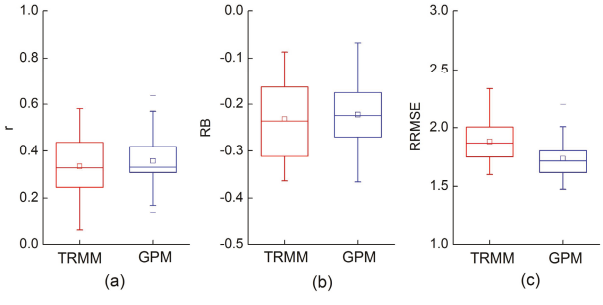


Figure 6. Evaluation results of TRMM 3B42 v7 and GPM IMERG v5 daily precipitation at individual gauge stations: (**a**–**c**) represent *r*, *RB* and *RRMSE*, respectively.

Figure 7 shows the spatial distribution of the evaluation results at each gauge station over the Huaihe River basin during the timeframe 2015–2017. Figure 7a shows the statistical results of mean annual rainfall at each gauge station, and Figure 7b,d represents the *r*, *RB*, and *RRMSE* of 3B42 v7 and IMERG v5 precipitation estimates, respectively. The lower *r* values were mainly distributed in the western and northern parts of the study area, where the terrain is relatively complex. The large absolute *RB* generally occurred at the gauge station with small mean annual precipitation. It could be because the rainy days at this gauge station were few, while the satellite precipitation was generally not zero. It was found that *r* of IMERG v5 was larger than the one of 3B42 v7 in the northwest of the Huaihe River basin while lower in the southeast; however, the absolute *RB* showed the contrary pattern. The *RRMSE* of 3B42 v7 was larger than that of IMERG v5 in all 43 stations.

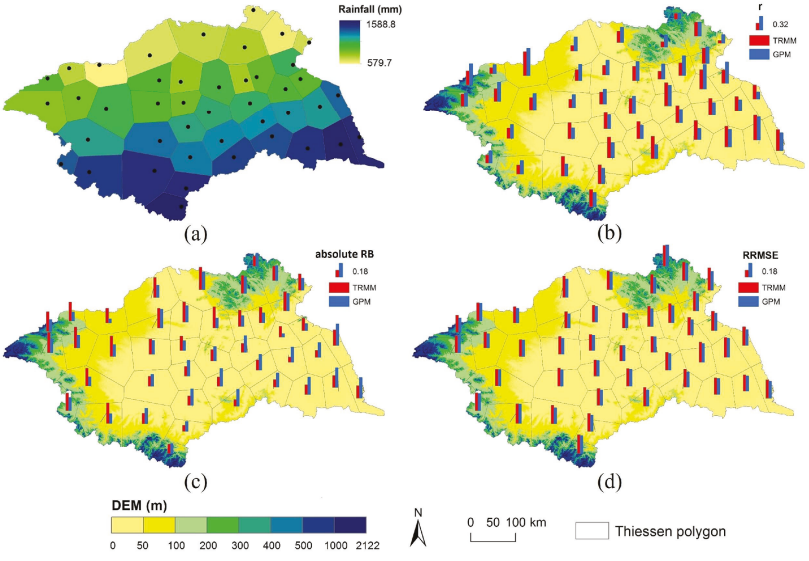


Figure 7. Evaluation results at each individual gauge station in the Huaihe River basin from 2015 to 2017. (**a**) Annual rainfall of rain gauge; (**b**–**d**) *r*, absolute *RB*, and *RRMSE* of 3B42 v7 and IMERG v5 precipitation products.

To further analyze the influence of elevation and mean annual precipitation on the *r*, absolute *RB* and *RRMSE* at each gauge station, a regression analysis was conducted, and the optimal fitting results are shown in Table 3. The elevation had a negative correlation with *r* (logarithmic regression). The *RRMSE* of IMERG v5 had a linear correlation with elevation (r = 0.38), and the *r* of 3B42 v7 had a linear correlation with mean annual rainfall.

Table 3. Regression results of three statistical metrics versus elevation and annual rainfall at all gauge stations.

	r					
	r		***Absolute RB***		***RRMSE***	
	TRMM	**GPM**	**TRMM**	**GPM**	**TRMM**	**GPM**
Elevation	$-0.40^{\log}$	$-0.42^{\log}$	#	#	#	0.38^{lin}
Annual rainfall	0.56^{lin}	#	$-0.54^{\exp}$	$0.36^{\exp}$	#	#

Notation: The symbol "#" means the regression results were not significant; Superscript represents the optimal fit, where "exp" means exponential regression, "lin" means linear regression, and "log" means logarithmic regression.

3.3. Performance of Satellite Precipitation Estimates at Different Rainfall Thresholds

Figure 8a,b presents the probability distribution function (PDF) and the mean daily precipitation estimated by gauge data and satellite products at different rainfall thresholds. The PDF and mean daily precipitation are of great importance for hydrological applications. The PDF of the two satellite precipitation products in Figure 8a were mainly concentrated at the rainfall rate of 0–0.5 mm/day (above 35%). The 3B42 v7 and IMERG v5 precipitation estimates had an obvious overestimation of PDF at small rainfall from 0–0.5 mm/day, while there was a tendency to underestimate the gauge precipitation when the rainfall rate was greater than 0.5 mm/day. The satellite precipitation estimates show similar mean daily precipitation to gauge observations in Figure 8b. The satellite estimates overestimated the mean daily precipitation at the rainfall rates from 0.5 to 25 mm/day, but tended to underestimate light (0–0.5 mm/day) and heavy (>25 mm/day) precipitation. Specifically, there was an obvious underestimation when the rainfall rates were greater than 100 mm/day.

The daily results of three metrics (*POD*, *CSI*, and *FAR*) for different rainfall thresholds are shown in Figure 9. High values of *POD* and *CSI* at rainfall rates from 0.5–2.0 mm/day indicated that satellite precipitation estimates had better performance. The *POD* and *CSI* were reduced, and the *FAR* was increased with increasing rainfall rate, which indicated the limited capability of satellite sensors for detecting heavy rainfall events. The relatively high *POD* and *CSI* also showed that IMERG v5 was better than 3B42 v7 for detecting precipitation in the study area, especially for rainfall rates of 0–2.0 mm/day.

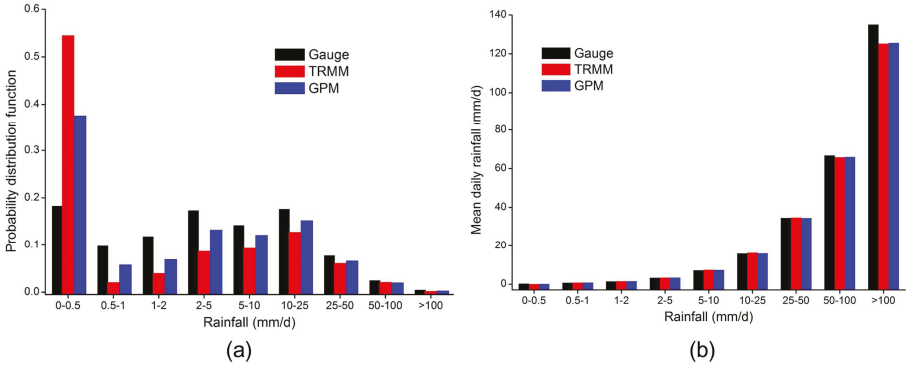


Figure 8. Probability distribution function (**a**) and mean daily precipitation (**b**) at different rainfall thresholds.

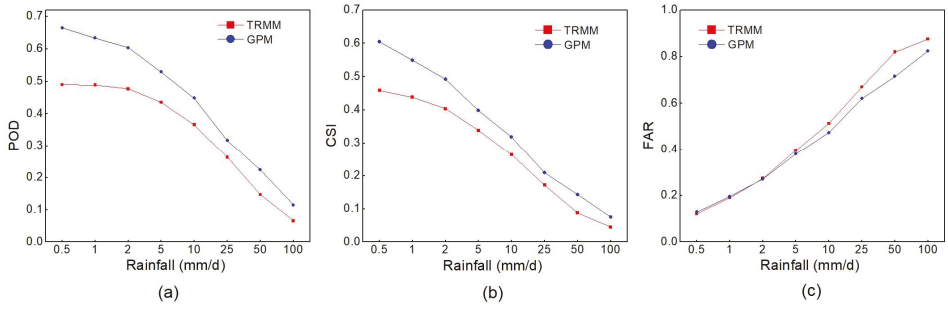


Figure 9. Daily results of three metrics for different rainfall thresholds: (**a**–**c**) represent *POD*, *CSI*, and *FAR*, respectively.

3.4. Influence of Latitude and Elevation on Satellite Precipitation

Figure 10 shows the continuous rainfall events of the rain gauge, 3B42 v7 and IMERG v5 precipitation data over the Huaihe River basin from 2015 to 2017. The daily rainfall rate was plotted according to its occurrence time and latitude. Both satellite products could better characterize the gauge precipitation distribution, while the relatively sparse gauge sites led to the loss of some important rainfall events, such as the rainfall events from 31°–34° N in late January and the strong precipitation in November. Satellite precipitation estimates can effectively detect spatial precipitation patterns and might be an alternative in the areas where the rain gauge stations are sparse.

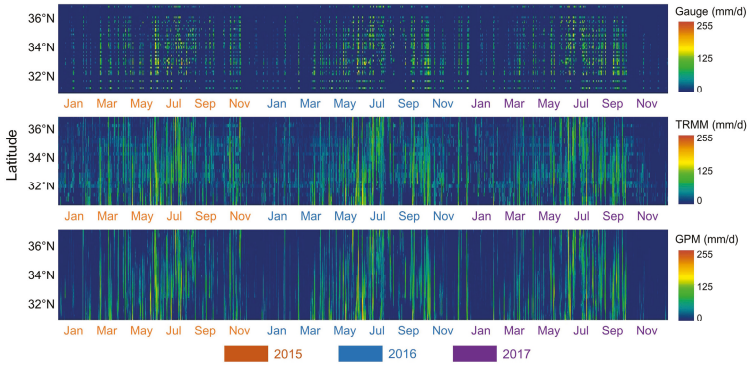


Figure 10. Comparison of satellite products (TRMM 3B42 v7 and GPM IMERG v5) and gauge observations with latitudes over the Huaihe River basin from 2015 to 2017.

Figure 11 presents the scatterplots of mean daily satellite precipitation against elevation over the Huaihe River basin. The points in Figure 11 with different colors representing the rainfall at different thresholds. The gray and black dots represent the mean daily satellite precipitation below and above the threshold rainfall rate of 3.0 mm/day, respectively. Various regression methods such as logarithm regression, exponential regression, and linear regression were tried to choose the best fitted regression between elevation and precipitation. The relationship between topography and precipitation was not significant over the whole basin. However, for rainfall rates less than 3.0 mm/day, the precipitation decreased logarithmically with elevation (the r values were 0.43 for 3B42 v7 and 0.45 for IMERG v5); for rainfall rates greater than 3.0 mm/day, the precipitation increased linearly with elevation (the r values were 0.55 for 3B42 v7 and 0.34 for IMERG v5).

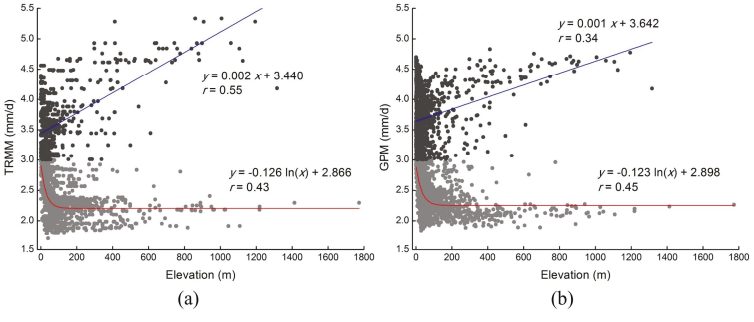


Figure 11. Scatterplots of mean daily precipitation of TRMM 3B42 v7 (**a**) and GPM IMERG v5 (**b**) precipitation products against the elevation over the Huaihe River basin.

4. Discussion

It should be highlighted that this study was the first to assess the performance of the latest released IMERG v5 precipitation product for three complete years of 2015–2017 over the climatic transition zone of the Huaihe River basin. It was found that both satellite precipitation products significantly underestimated the gauge precipitation (*RB* was –0.273 for 3B42 v7 and –0.176 for IMERG v5) and had relatively small *r* when the gauge stations with zero rainfall were eliminated in Figure 2a, which was mainly due to the absence of the whole rainfall process record of the satellite sensors (the temporal resolutions for IMERG v5 and 3B42 v7 are half-hourly and 3-hourly) leading to the inaccurate monitor of the smaller rainfall events. Usually, the TRMM precipitation radar (PR), whose emission wavelengths are relatively long, has a minimum measurable rain rate as low as 0.7 mm/h [39]. While a key advancement of GPM is the extended capability to measure light rain (less than 0.5 mm/h), solid precipitation, and the microphysical properties of precipitating particles compared to TRMM; the GPM Dual-frequency Precipitation Radar (DPR) is more sensitive to light rain rates and snowfall [10]. Additionally, simultaneous measurement by the overlapping of Ka/Ku-bands of the DPR provides new information on particle drop size distributions over moderate precipitation intensities. Thus, the absolute *RB* of IMERG v5 shown in Figure 2a was relatively lower than that of 3B42 v7 at daily scale.

The GPM IMERG v5 and TRMM 3B42 v7 precipitation products overestimated precipitation at monthly and annual scales (Figure 2b,c), which was mainly because the satellite precipitation estimates were usually greater than zero when the gauges showed no rainfall. When the time scale increased, the *r* values were significantly improved, and the *RRMSE* values were significantly reduced. Similar findings have been reported in previous evaluation results [15,22]. A reason for the higher accuracy of satellite precipitation products at monthly and annual scales was that the GPCC monthly product was used for bias-correction of the 3B42 v7 and IMERG v5 products [11,40].

A solid understanding of the distribution of daily precipitation is of great importance in hydrological applications. PDF emphasizes lighter rainfall rates, making it more suitable for evaluating the ability of satellite precipitation to detect light precipitation [14]. The PDF of satellite precipitation shown in Figure 8a was consistent with that of gauge observations for the rainfall rate greater than 0.5 mm/day and tended to underestimate the gauge precipitation in this study. An obvious overestimation of satellite precipitation frequency occurred during light rainfall from 0–0.5 mm/day. Yang and Geng [23] also reported that the TRMM-era precipitation estimates tended to underestimate light rainfall. The precipitation distribution in Figure 8b showed an underestimation of satellite estimates when the rainfall rates were either less than 0.5 mm/day or greater than 25 mm/day. Ebrahimi et al. [41] pointed out that satellite sensors are not sensitive to light rainfall and heavy rainfall, which might lead to unsatisfactory monitoring of extreme rainfall. The IMERG v5 product, with relatively high *POD* and *CSI*, showed better performance in precipitation detection of the Huaihe River basin. This was mainly because the GPM combined Instrument (GMI) sensor can capture light precipitation better than the TRMM combined Instrument (TMI) [42]. Additionally, Figure 9 demonstrated the *FAR* was increased, but the *POD* and *CSI* were reduced with increasing rainfall rate for the two satellite precipitation products, suggesting the limited capability of satellite sensors in detecting intense rainfall events [24]. Conversely, the high intensity rainfall makes radar signals attenuate, so the strong rainfall cannot be accurately measured [43]. Then again, because the satellite precipitation is the homogenization data of the grid downscaling, it is unable to reflect the local heavy rainfall characteristics.

It is widely acknowledged that topography has a strong effect on precipitation [44–46], and the relationship between altitude and precipitation varies in different regions, especially in the climatic transition region. During this study, the negative logarithmic relationship between elevation and mean daily rainfall was found when the rainfall rate was less than 3.0 mm/day, which could be due to the effect of terrain and latitude [47]. The rainfall increased linearly with topography when the rainfall rates were above 3.0 mm/day. Regions where the precipitation rates were above 3.0 mm/day were generally located in the south of the Huaihe River basin (also to the south of 35° N, see Figures 1 and 5). This could be attributed to the effects of the western Pacific subtropical high (abbreviated as subtropical high),

which is one of the main atmospheric circulation systems influencing the weather and climate change in China [48]. The southward–northward shift of the precipitation belt in China is consistent with the seasonal activities of the subtropical high, and rainfall is usually located at 5°–8° latitude north of the ridge position of the subtropical high [49,50]. During summer, the ridge of the subtropical high is located at about 25° N. The southern part of the Huaihe River basin is under the subtropical high influence, and such influence decreases with latitude. Since the elevation in the south of the Huaihe River basin also decreases with latitude, precipitation increases with elevation when the precipitation rates were above 3.0 mm/day. The different regression relationships between elevation and precipitation were found and the results were discussed from the perspective of atmospheric circulation, which could provide the reference for the study of rainfall in other climatic transition regions.

IMERG v5 had larger *r* and smaller *RRMSE* than 3B42 v7 on multiple temporal scales. The relatively high *POD* and *CSI* of IMERG v5 also indicated its higher detection ability for daily rainfall, especially for rainfall rates of 0–2.0 mm/day. Daily precipitation or even hourly precipitation is very important input for hydrological models. The lack of dense rain gauges at the basin scale leads to great uncertainty in practical hydrological simulations [51]. IMERG v5, which has high temporal resolution and accuracy, might have better potential in future hydrological applications. The GPM precipitation with higher spatial resolution (0.1°) showed a better ability for precipitation observation in small areas compared to the coarse resolution (0.25°) of the TRMM products on a spatial scale. Additionally, IMERG v5 can make up for the limitations of 3B42 v7 regarding detecting rainfall over water bodies due to the employment of a unified and updated PMW retrieval algorithm [52]. It was demonstrated that IMERG v5 was more suitable for characterizing the spatial variation of precipitation. However, it is worth pointing out that the accuracy of satellite precipitation at daily scale is still poor, which limits the application of satellite precipitation products to some extent [41]. Xu et al. [40] found that the performance of GPM and TRMM strongly depends on topography and rainfall intensity over the southern Tibetan Plateau. Since the accuracy of satellite precipitation products usually is affected by the rainfall rate, latitude, and elevation in the climate transition zone, fusion of multi-source precipitation products (satellite, radar and measured data) and geographic background information (terrain and geographic location) to enhance the spatiotemporal accuracy of IMERG v5 data will be the hot spot of future study.

It should be noted that the 3B42 v7 data in this study were resampled to the same spatial resolution (0.1°) as IMERG v5 data by using the standard bilinear interpolation method to make them comparable [13,15]. However, potential errors would possibly be introduced during the resampling procedure. To investigate the impact of the interpolation method, the error analysis of the initial TRMM 3B42 v7 data (0.25°) was conducted and the evaluation results were shown in Table 2. It was found that the GPM IMERG v5 (0.1°) as a whole had the best performance. The initial (0.25°) and the interpolated (0.1°) TRMM data had different performances at multiple scales, which demonstrated that the comparison of different interpolation methods is necessary to obtain higher precision for precipitation. Moreover, the nearest pixel to a rain gauge was extracted to conduct the comparison between the point-based rain gauge and pixel-based GPM or TRMM in the study. However, there is a typical scale mismatch issue between point-based rain gauge data and the gridded precipitation products [22]. Many interpolation methods (inverse distance weighting, kriging, spline, and Thiessen polygon for example) [22,33,53] and grid-to-point techniques (such as nearest neighbor, bilinear weighted interpolation) [41] have been proposed to compare the point-based rain gauge data and pixel-based satellite precipitation data. Each method has its advantages and disadvantages, and its performance depends on various factors and varies from region to region [22]. It is an interesting topic to develop a common method in all study areas for comparing the point-based rain gauge and pixel-based satellite precipitation.

5. Conclusions

An early and comprehensive assessment of the latest released GPM IMERG v5 and TRMM 3B42 v7 precipitation products was carried out over the Huaihe River basin with a climate transition. The accuracy of satellite precipitation retrievals was typically influenced by the rainfall rate and

local topographical features in the climate transition zone. Both satellite precipitation products agreed with the gauge observations at monthly and annual scales, providing effective data source for hydrometeorology research to supplement sparse gauge observations. However, significant overestimation of daily precipitation at each individual gauge station was found, indicating the necessity for evaluation before application with certainty. The initial assessment on detecting rainfall event frequency at different rainfall thresholds showed that both products overestimated precipitation distribution at the rainfall rates 0.5 to 25 mm/day, but tended to underestimate light (0–0.5 mm/day) and heavy (>25 mm/day) precipitation (especially for the torrential rain, above 100 mm/day) in the basin. The latest released IMERG v5 product had better performance for detecting precipitation and provided more accurate precipitation estimates than TRMM 3B42 v7 data due to finer spatial and temporal resolutions. Therefore, there is an urgent need to assess the potential of the GPM IMERG v5 product in hydrological application in a range of different environments.

Author Contributions: C.C. analyzed the data and wrote the draft of the manuscript; C.C. and Z.D. performed the experiments; Z.L., J.Z. and G.T. gave comments and revised the manuscript; K.M. handled the gauge data. Q.C. conceived of this study, gave comments and significantly revised the manuscript.

Funding: This research was supported by the National Key Research and Development Program of China (2016YFC0502205, 2016YFC0400902), the National Natural Science Foundation of China (51709179) and the fund of Nanjing Hydraulic Research Institute (Y918011).

Acknowledgments: We are grateful to the providers of operational satellite precipitation products and rain gauge measurements for making data available to us.

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

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. Duan, Z.; Bastiaanssen, W.G.M. 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]
3. 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]
4. Allen, M.R.; Ingram, W.J.; Stainforth, D.A. Constraints on future changes in climate and the hydrologic cycle. *Nature* **2002**, *419*, 224. [CrossRef] [PubMed]
5. Sorooshian, S.; Hsu, K.L.; Gao, X.; Gupta, H.V.; Imam, B.; Dan, B. Evaluation of persiann system satellite–based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2035–2046. [CrossRef]
6. 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]
7. Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Kachi, M. 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]
8. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Michaelsen, J. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 150066. [CrossRef] [PubMed]
9. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.; 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]
10. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Iguchi, T. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 701–722. [CrossRef]
11. 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). In *Algorithm Theoretical Basis Document (ATBD)*; NASA/GSFC: Greenbelt, MD, USA, 2014.

12. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Integrated Multi-SatellitE Retrievals for GPM (IMERG) Technical Documentation*; NASA/GSFC Code; NASA/GSFC: Greenbelt, MD, USA, 2015.
13. 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]
14. 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]
15. 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]
16. Gaona, M.R.; Overeem, A.; Leijnse, H.; Uijlenhoet, R. First-Year Evaluation of GPM Rainfall over The Netherlands: IMERG Day 1 Final Run (V03D). *J. Hydrometeorol.* **2016**, *17*, 2799–2814. [CrossRef]
17. Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Liu, Z.; Norouzi, H.; Pai, D.S. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.* **2016**, *556*, 865–875. [CrossRef]
18. 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]
19. Asong, Z.E.; Razavi, S.; Wheater, H.S.; Wong, J.S. Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over Southern Canada against Ground Precipitation Observations: A Preliminary Assessment. *J. Hydrometeorol.* **2017**, *18*, 1033–1050. [CrossRef]
20. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM Precipitation Products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]
21. Li, Z.; Yang, D.; Gao, B.; Jiao, Y.; Hong, Y.; Xu, T. Multiscale hydrologic applications of the latest satellite precipitation products in the Yangtze River Basin using a distributed hydrologic model. *J. Hydrometeorol.* **2015**, *16*, 407–426. [CrossRef]
22. 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]
23. Yang, X.; Geng, W. Accuracy evaluation of TRMM-based Multi-satellite Precipitation in Huai River Basin. *Water Resour. Power.* **2016**, *34*, 1–5.
24. Sun, R.; Yuan, H.; Liu, X.; Jiang, X. Evaluation of the latest satellite–gauge precipitation products and their hydrologic applications over the Huaihe River basin. *J. Hydrol.* **2016**, *536*, 302–319. [CrossRef]
25. Stocker, E.F.; Kelley, O. GPM V05 Gridded Text Products. Available online: https://ntrs.nasa.gov/search.jsp?R=20170009623 (accessed on 2 December 2017).
26. Zulkafli, Z.; Buytaert, W.; Onof, C.; Manz, B.; Tarnavsky, E.; Lavado, W.; Guyot, J.L. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean–Amazon river basins. *J. Hydrometeorol.* **2014**, *15*, 581–592. [CrossRef]
27. Huffman, G.J.; Bolvin, D.T. TRMM and Other Data Precipitation Data Set Documentation, Mesoscale Atmospheric Processes Laboratory, NASA Global Change Master Directory Doc. Available online: https://pmm.nasa.gov/sites/default/files/document_files/3B42_3B43_doc_V7.pdf (accessed on 8 April 2015).
28. Sandwell, D.T. Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. *Geophys. Res. Lett.* **1987**, *14*, 139–142. [CrossRef]
29. 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]
30. Duan, L.; Fan, K.; Li, W.; Liu, T. Spatial downscaling algorithm of TRMM precipitation based on multiple high-resolution satellite data for Inner Mongolia, China. *Theor. Appl. Climatol.* **2017**, 1–15. [CrossRef]
31. Zhang, Q.; Singh, V.P.; Li, J.; Chen, X. Analysis of the periods of maximum consecutive wet days in China. *J. Geophys. Res. Atmos.* **2011**, *116*. [CrossRef]
32. Xu, W.; Li, Q.; Wang, X.L.; Yang, S.; Cao, L.; Feng, Y. Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices. *J. Geophys. Res. Atmos.* **2013**, *118*, 9708–9720. [CrossRef]

33. Yong, B.; Ren, L.L.; Hong, Y.; Wang, J.H.; Gourley, J.J.; Jiang, S.H.; 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]
34. 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]
35. 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]
36. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*; Academic Press: Cambridge, MA, USA, 2011.
37. 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]
38. 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. *Global Planet. Chang.* **2015**, *133*, 188–200. [CrossRef]
39. Kawanishi, T.; Takamatsu, H.; Kozu, T.; Okamoto, K.; Kumagai, H. TRMM precipitation radar. *Adv. Space Res.* **2000**, *25*, 969–972. [CrossRef]
40. 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]
41. Ebrahimi, S.; Chen, C.; Chen, Q.; Zhang, Y.; Ma, N.; Zaman, Q. Effects of temporal scales and space mismatches on the TRMM 3B42 v7 precipitation product in a remote mountainous area. *Hydrol. Process.* **2017**, *31*. [CrossRef]
42. 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]
43. Cheema, M.; Bastiaanssen, W. Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *Int. J. Remote Sens.* **2012**, *33*, 2603–2627. [CrossRef]
44. Bonacina, L.C.W.; Poulter, R.; Ashmore, S.; Manley, G.; Bonacina, L. Orographic rainfall and its place in the hydrology of the globe. *Q. J. Roy. Meteor. Soc.* **1945**, *71*, 41–55.
45. Hughes, M.; Hall, A.; Fovell, R.G. Blocking in areas of complex topography, and its influence on rainfall distribution. *J. Atmos. Sci.* **2009**, *66*, 508–518. [CrossRef]
46. Kumari, M.; Singh, C.K.; Bakimchandra, O.; Basistha, A. Geographically weighted regression based quantification of rainfall–topography relationship and rainfall gradient in Central Himalayas. *Int. J. Climatol.* **2017**, *37*, 1299–1309. [CrossRef]
47. Qing, Y.; Zhu-Guo, M.A.; Liang, C. A preliminary analysis of the relationship between precipitation variation trends and altitude in China. *Atmos. Ocean. Sci. Lett.* **2011**, *4*, 41–46. [CrossRef]
48. Qu, W.; Wang, J.; Gao, S.; Wu, T. Effect of the strengthened western Pacific subtropical high on summer visibility decrease over eastern China since 1973. *J. Geophys. Res. Atmos.* **2013**, *118*, 7142–7156. [CrossRef]
49. Ueda, H.; Ohba, M.; Xie, S.P. Important factors for the development of the Asian–Northwest Pacific summer monsoon. *J. Clim.* **2009**, *22*, 649–669. [CrossRef]
50. Xie, S.P.; Hu, K.; Hafner, J.; Tokinaga, H.; Du, Y.; Huang, G.; Sampe, T. Indian Ocean capacitor effect on Indo–Western Pacific climate during the summer following El Niño. *J. Clim.* **2009**, *22*, 730–747. [CrossRef]
51. Meng, J.; Li, L.; Hao, Z.; Wang, J.; Shao, Q. Suitability of TRMM satellite rainfall in driving a distributed hydrological model in the source region of Yellow River. *J. Hydrol.* **2014**, *509*, 320–332. [CrossRef]
52. Tang, G.; Long, D.; Hong, Y. Systematic anomalies over inland water bodies of High Mountain Asia in TRMM precipitation estimates: No longer a problem for the GPM era? *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1762–1766. [CrossRef]
53. Awange, J.; Forootan, E. An evaluation of high-resolution gridded precipitation products over Bhutan (1998–2012). *Int. J. Climatol.* **2016**, *36*, 1067–1087.

remote sensing

MDPI

Article

Rain Microstructure Parameters Vary with Large-Scale Weather Conditions in Lausanne, Switzerland

Wael Ghada [1,*], Allan Buras [1], Marvin Lüpke [1], Christian Schunk [1] and Annette Menzel [1,2]

[1] Department of Ecology and Ecosystem Management, Technical University of Munich, Hans-Carl-von-Carlowitz-Platz 2, D-85354 Freising, Germany; buras@wzw.tum.de (A.B.); luepke@wzw.tum.de (M.L.); schunk@wzw.tum.de (C.S.); amenzel@wzw.tum.de (A.M.)

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

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

Received: 9 April 2018; Accepted: 22 May 2018; Published: 23 May 2018

Abstract: Rain properties vary spatially and temporally for several reasons. In particular, rain types (convective and stratiform) affect the rain drop size distribution (DSD). It has also been established that local weather conditions are influenced by large-scale circulations. However, the effect of these circulations on rain microstructures has not been sufficiently addressed. Based on DSD measurements from 16 disdrometers located in Lausanne, Switzerland, we present evidence that rain DSD differs among general weather patterns (GWLs). GWLs were successfully linked to significant variations in the rain microstructure characterized by the most important rain properties: rain intensity (R), mass weighted rain drop diameter (D_m), and rain drop concentration (N), as well as $Z = AR^b$ parameters. Our results highlight the potential to improve radar-based estimations of rain intensity, which is crucial for several hydrological and environmental applications.

Keywords: synoptic weather types; drop size distribution (DSD); microstructure of rain; disdrometer; radar reflectivity–rain rate relationship

1. Introduction

Ground measurements of the rain drop size distribution (DSD) have attracted scientists' attention for a variety of reasons. (1) Variations in the rain DSD influence the relationship between the radar reflectivity factor (Z) and the rainfall intensity (R) [1]. An accurate estimation of R is essential for important applications in hydrology, water management, and flood protection; (2) Rain kinetic energy depends on rain DSD and influences soil erosion in different scales [2,3]; (3) Rain drops have the potential to wash particulate pollutants off impermeable surfaces, affecting the ecology of urban waterways [4]; (4) Rainfall distorts telecommunication signals and radio propagation, especially at high frequencies [5], all of which are influenced by rain DSD [6]; (5) Specific methods used to separate rain type (convective and stratiform) are based on rain DSD [7,8]. Consequently, different tools and methods have been developed to detect rain DSD and to improve our understanding of its underlying processes.

Automated disdrometers measure rain DSD on the ground with high temporal resolution. These measurements are used to detect variations in the rain DSD on different temporal and local scales. Rain DSD parameters may vary at a scale of a few meters, and the small sampling areas of disdrometers contribute to this variation [9]. The rain DSD also varies within the area of one weather radar pixel (1 km^2) [10]. The detected spatial variation increases with the increasing distance between disdrometers [11]. It exists even when decreasing the temporal resolution [12]. DSD parameter variation has also been confirmed on regional [13] and global [14] scales. Seasonality, rain type, climate

regime, and their combinations also result in different rain DSD [15–17]. The rain DSD even differs from event to event [18].

Rain DSD variation is linked to the microphysical processes of rain formation such as coalescence [19], fragmentation [19,20], evaporation [19,21,22], and size sorting of rain drops by wind drafts [23]. Also conditions such as the height of freezing levels, the boundary layer relative humidity [24], and aerosols' abundance and properties [25,26] influence rain DSD.

Rain DSD evolves under the complex micro- and mesoscale conditions. For example, topography influences the interaction between rain forming processes in orographic rain [27–29]. In addition, rain DSD varies within the stages of squall lines [16,30]. It also varies in convective versus stratiform rain, continental versus maritime rain, and warm versus cold rain [24,31,32]. Munchak et al. [24] explained that convective rain is associated with strong updrafts where cloud tops are above the freezing level, and both warm and cold rain forming processes occur. In contrast, updraft in stratiform rain is limited, and cold processes prevail. Rain DSD observations confirm the distinct microstructure between both rain types, with larger drops and higher intensities in convective rain [8,13,33,34]. Since they are based on prevailing flow directions and source regions of air masses govern several of the aforementioned influencing factors at the local scale, we hypothesize that large scale weather conditions may help to improve our knowledge about and thus modeling of parameters related to rain DSD.

Weather type classification has proven to be useful in understanding climatic processes and the link between the variability of local weather conditions and atmospheric circulations [35,36]. Weather types represent the atmospheric conditions that prevail over a large area and remain relatively stable for several days (a minimum of three days) [37]. These atmospheric conditions include the location of sea-level, semi-permanent pressure centers, the position and paths of frontal zones and the existence of cyclonic or anticyclonic circulation types [38]. Methods have been developed and used to classify weather types into patterns and groups [35]. In this paper, the classification of weather types by Hess and Brezowsky [39] is used. Their classes are called general weather patterns and are abbreviated as GWL, which is derived from the German word "Großwetterlage." In this classification, each day is assigned to one of 29 GWL classes or a transitional type, and each GWL lasts three days or more [40]. The relationship between large-scale weather types and local weather phenomena has been verified previously, e.g., some weather types have a major influence on the generation of extremely hot days [41]. Moreover, GWLs influence stream flows [42], floods [43,44], forest fires [45], and even debris-flow events [46]. More directly, GWLs have a substantial influence on precipitation totals, particularly on their extreme values [47–49] and even seasonal extremes [50]. Consequently, because changes in weather type frequencies relate to changes in precipitation amounts and extremes [51], a better understanding of the impact of large-scale circulations on intense (sub-)daily precipitation is required [52].

Synoptic weather types provide a combination of meteorological factors that influence the rain DSD. Fernandez-Raga et al. [53] provided convincing evidence that weather patterns with northern components are associated with smaller rain drops in Leon, Spain. On the other hand, western and southwestern components carry humidity from the Atlantic and have larger drops. Hachani et al. [29] verified the different rain DSD characteristics and Z–R relationship between different factors including weather patterns, seasons, location, and rain fall type in the Cévennes–Vivarais region, France. They also investigated the combined effects of weather patterns and locations. However, the combined effect of rain type and weather patterns on the rain DSD has not yet been addressed.

Based on these outlined knowledge gaps, we address two general questions here.

(I) Does the rain microstructure differ between general weather patterns in both convective and stratiform rain types?
(II) How much does this influence the radar reflectivity and rain rate, i.e., the Z–R relation?

2. Materials and Methods

2.1. Data Sources

We based our analyses on a published dataset collected by 16 "OTT PARSIVEL" disdrometers spanning the period from December 2008 through September 2010 in Lausanne, Switzerland (the EPFL-LTE campaign) [54]. The disdrometers were deployed over an area of less than 1 km^2. The aim of the original campaign was to assess the spatial variability within the area covered by one pixel of weather radar. The dataset has temporal resolutions of 20 s and 30 s. A detailed map of the disdrometers locations, a description of the campaign, data structure, and the measured parameters is given in Jaffrain et al. [10,11]. For the purpose of this paper, when comparing properties between GWLs, all the intervals were aggregated to a one-minute temporal resolution. Each minute interval measured by a device was considered to be a realization of the rain properties of the associated GWL, which means that up to 16 simultaneous realizations of rain properties were available for each minute of rain. It is important to note that the 16 disdrometers were not deployed at the same time, and not all of them measured for the entire 22-month period. This resulted in variations in the total rain duration and amount measured across the devices. The accumulated rain amount and duration measured by each device along with the start and end dates of the measurements are provided in Table S1.

There was also a difference between the total mean amount of rain recorded over the network that is reported in the published dataset (1275 mm) [54] and the total mean reported in our analysis (1027 mm). This is a result of the filtering procedures, which are explained in Section 2.5.

The German Meteorological Service (DWD) performs a daily classification of the general weather patterns. This classification is based on Hess and Brezowsky [39] and includes 29 GWL classes and 1 transitional class. The daily records of this classification are published online [55–57]. The 29 GWLs can be aggregated according to the patterns of the air parcel movements (zonal, meridional, or mixed), the rotation direction (cyclonic or anti-cyclonic), and the flow direction. Table 1 lists all the GWLs and the corresponding five major classes of flow directions [50]. Each GWL is associated to one of these five classes, which indicates the origin of the moving air parcels, as derived from the wind flow direction at upper levels (500 hPa). These five classes are referred to as the flow directions throughout the text.

2.2. Measurement Device

The "OTT PARSIVEL" disdrometer is an optical, laser-based device that measures the size and velocity of hydrometeors. It counts the number of hydrometeors that fall in 32 diameter ranges and 32 velocity ranges. The first two diameter ranges are not used by the disdrometer and are always empty leaving an actual detection range of 0.31–24.5 mm. The 32 velocity ranges extend from 0.05 $m{\cdot}s^{-1}$ to 20.8 $m{\cdot}s^{-1}$. A list of the diameter ranges is available in [58]. The disdrometer uses a horizontal laser sheet with a detection area of 54 cm^2 (180 mm × 30 mm). Whenever a hydrometeor passes through this sheet, the laser signal is reduced and the diameter is calculated according to this reduction. The velocity is derived from the size of the hydrometeor and the time it takes to pass the laser detection sheet. In addition, OTT PARSIVEL disdrometers provide an automated separation for the precipitation type based on the combination of the size and velocity values in each interval. A more detailed description of the OTT PARSIVEL device is provided by Löffler-Mang and Joss [59].

The disdrometer raw output was used in this study to obtain the values of the rain intensity R [$mm{\cdot}h^{-1}$], reflectivity Z [$mm^6{\cdot}m^{-3}$ and dBZ], total drop concentration N [m^{-3}], and mass weighted diameter D_m [mm]. Equations (A1)–(A5) [16,60,61] were used to obtain these values (Appendix A). The parameters, R, D_m, and N, best illustrate the distinct physical rain properties for both rain types.

2.3. Separation of Convective and Stratiform Rain Intervals

Several methods to separate convective and stratiform rain are available in the literature. These methods are based on instrumental data, e.g., from radar and wind profilers. Some methods have also suggested the use of rain DSD parameters [7,8]. Here the separation of convective and stratiform rain intervals follows the simple procedure proposed by Bringi et al. [15], which has already been adopted in several

studies [18,62]. Rain is regarded as convective if R is larger than 10 mm·h^{-1} or if the standard deviation of the rain intensities over five consecutive two-minute intervals is larger than 1.5 mm·h^{-1}; otherwise, the rain interval is considered to be stratiform.

Table 1. General weather patterns (GWLs) according to Hess and Brezowsky [39] and their classification based on the flow direction [50]. This table also includes the accumulated rain duration over all devices for each GWL in both rain types. Gray cells indicate GWLs with small sample sizes, which are not further considered in the analyses.

Flow Direction	GWL		Accumulated Duration [min]	
			Convective	Stratiform
W: Westerly	WA	Anticyclonic Westerly	2	441
	WZ	Cyclonic Westerly	12,101	79,492
	WS	South-Shifted Westerly	1288	35,715
	WW	Maritime Westerly	3454	30,797
S: Southerly	SWA	Anticyclonic Southwesterly	623	4404
	SA	Anticyclonic Southerly	0	0
	SEA	Anticyclonic Southeasterly	9	4954
	SWZ	Cyclonic Southwesterly	15,129	83,369
	SZ	Cyclonic Southerly	5	296
	TB	Low over the British Isles	1001	7703
	TRW	Trough over Western Europe	8979	137,716
	SEZ	Cyclonic Southeasterly	37	11,670
NW–N: Northwesterly, Northerly	NWA	Anticyclonic Northwesterly	0	121
	NA	Anticyclonic Northerly	0	0
	HNA	Icelandic High, Ridge C.E. *	0	0
	HB	High over the British Isles	1	883
	NWZ	Cyclonic Northwesterly	1714	20,052
	NZ	Cyclonic Northerly	1507	4573
	HNZ	Icelandic High, Trough C.E.	1076	13,672
	TRM	Trough over C.E.	5967	32,031
CE: Main high/low pressure area over Central Europe (no specific prevailing direction)	HM	High over C.E.	2673	6464
	TM	Low (Cut-Off) over C.E.	2157	13,966
	BM	Ridge across C.E.	3358	18,977
NE–E: Northeasterly, Easterly	NEA	Anticyclonic Northeasterly	0	977
	HFA	Scandinavian High, Ridge C.E.	696	7734
	HNFA	High Scandinavia-Iceland, Ridge C.E.	0	0
	NEZ	Cyclonic Northeasterly	129	3711
	HFZ	Scandinavian High, Trough C.E.	0	6756
	HNFZ	High Scandinavia-Iceland, Trough C.E.	924	9361
-	U	Transitional	0	0

* Central Europe (C.E.).

2.4. Data Filtering

Error sources and shortcomings of disdrometers in general and OTT PARSIVEL disdrometers in particular include: wind effects, unstable laser source, splashing effect, drops falling through the edges of the detection area (margin fallers), non-rain particles such as insects and flying seeds, drops falling simultaneously through the device detection area (masking effect), and the small sampling area [9,63,64]. To minimize these errors, we adopted the procedure suggested and explained in detail by Friedrich et al. [64], composed of the following steps in the presented order.

1. Intervals characterized by high wind speed were completely removed from the dataset. These intervals were indicated by the existence of large drops D > 5 mm with a low falling speed V < 1 m·s^{-1}.
2. All particles larger than 8 mm in diameter were not considered to be rain and filtered out.
3. Particles with velocities higher than 160% of the associated terminal velocity for the diameter class center as given by Atlas et al. [65] were filtered out (margin fallers).
4. Small particles (D < 2 mm) with low velocity (<40% of the terminal velocity) were filtered out (splashing).
5. Large particles (D > 5 mm) with low velocity (<40% of the terminal velocity) were filtered out (hail).
6. For particles with a diameter between 2 mm and 5 mm, a fall velocity threshold was established based on the average terminal velocity of rain and graupel for each diameter bin. Then, all particles with a fall velocity below this threshold were filtered out.

In addition, we excluded:

1. Intervals associated with a damaged laser signal as reported in the raw output of the disdrometers;
2. Intervals associated with snow, hail, frozen rain, or mixed precipitation as reported by the disdrometer;
3. Intervals with rain intensities lower than 0.1 $mm \cdot h^{-1}$;
4. Intervals where three or less diameter bins recorded rain drops to ensure the existence of a drop size distribution; and
5. Intervals of rain that occurred in the transitional GWL U.

In total, the dataset was composed of rain DSD records in 25 GWLs, after excluding the transitional GWL U, and four other GWLs for which no rain was recorded (Table 1). These 25 GWLs form the base of the results in Section 3.1. However, when discriminating the rain type into convective and stratiform, the inhomogeneous frequency of GWLs, as well as the variation in the rain duration and rain amount per GWL, resulted in a varying number of rain intervals detected for each GWL–rain type combination. Some GWLs had only a small number of rain intervals per rain type. Therefore, to compare the rain microstructure (Section 3.2) and the Z–R parameters (Section 3.3), GWLs that had less than 500 min of accumulated convective duration or 1000 min of accumulated stratiform rain over all 16 devices (shaded cells in Table 1) were excluded from the analysis in these sections. This reduced the number of GWLs to 20 for stratiform rain and 16 for convective rain. The lower threshold in the case of convective rain was based on the generally lower duration of detected convective rain per GWL.

2.5. Assessing the Performance of the Z–R Relation

Rain intensity R [$mm \cdot h^{-1}$] retrieval from weather radars is generally based on the Z relation:

$$Z = A \times R^b, \tag{1}$$

where Z [$mm^6 \cdot m^{-3}$] is usually measured directly by the weather radar, while A and b are determined for each rain type separately. Then, Equation (1) is used to determine R. Here, conversely, the values of R and Z were calculated for each interval from the raw output of the disdrometers, as explained in Appendix A. Then, a linear regression of log(Z) and R was used to determine the values of A and b using two models. The general Z–R model takes only rain type into account, resulting in one value of A and one value of b for each rain type, convective and stratiform. In the GWL Z–R model, A and b are specifically determined for each combination of GWL and rain type.

To evaluate the performance of the two models, the mean absolute error (MAE) of the rain intensity (R) estimations of each model was used as recommended by Willmott and Matsuura [66]. The R values based on Equation (A1) were considered to be the accurate observed values, while the model outputs of R were considered to be the prediction. MAE is given by

$$\mathrm{MAE} = n^{-1} \sum_1^n |e_i|, \tag{2}$$

where n is the number of observations, e_i is the model prediction error: $e_i = Vm_i - Vo_i$, Vm_i is the model prediction value, Vo_i is the observed value, and i is the observation index.

The relative error was also used to assess the overestimation (or underestimation) of the total rain amount by both models for the entire dataset, as well as for each GWL-rain type combination:

$$\mathrm{RE} = \sum e_i / \sum Vo_i. \tag{3}$$

To obtain a robust evaluation of both models, we separated the dataset into a training subset (2/3 of the rain intervals in each GWL/rain type combination chosen randomly) and a testing subset (the remaining 1/3 of the dataset). Both models were built based on the training subset, then evaluated by calculating MAE and RE from the testing subset. The whole process was repeated 100 times and the reported values in Section 3.3 represent the mean of these repetitions.

The 16 disdrometers were deployed within 1 km^2, which allows comparisons of the spatial variation of A and b at this scale versus their GWL-specific variations. Therefore, the A and b values were determined for each device-rain type combination and then were compared with the A and b values of the GWL Z–R model.

2.6. Statistical Tests

For the parameters R, D_m, and N, a normality check was performed by means of visual inspections (QQ plots and histograms) and a Shapiro–Wilk normality test. Because all parameters were non-normally distributed (p-value of Shapiro–Wilk test < 0.05), non-parametric tests were chosen. To compare more than two groups, the Kruskal–Wallis test was performed to detect the existence of significantly different parameters (p-value < 0.05) between GWLs and a pairwise Wilcoxon test was performed to check the significance of the differences between each pair of GWLs. These steps were performed for each rain type separately. A similar procedure was followed when comparing these parameters between flow directions (cf. Table 1).

To compare the spatial variability of the Z–R parameters (A and b) against their GWL variability, the A and b values were calculated for each device and then for each GWL. A normality check was performed for the A and b values. An F-test was used for the normally distributed parameters (only A in the case of stratiform rain), and Levene's test was used for the non-normally distributed parameters. The normality check and the variance comparisons were performed separately for each rain type.

Data filtering, analyses, and plots were performed using the statistical computing environment R [67], the integrated development environment RStudio [68], and the data visualization package ggplot2 [69].

3. Results

3.1. Variation in Rain Duration, Amount and Rates with Rain Type and GWL

This section provides an overview of the rain records in general and within each rain type and GWL. We aim at demonstrating the variation in convective rain proportions between GWLs, which emphasizes the importance of separation by rain type for comparing rain microstructure among GWLs (see Section 4.2).

After filtering, a total of 598,665 one-minute DSD samples (42% of the total dataset) were included in the analyses. The total detected rain amount was 16,439 mm accumulated over the 16 disdrometers, with a median rain intensity of 0.79 $mm \cdot h^{-1}$ (standard deviation (sd): 3.37 $mm \cdot h^{-1}$). Convective rain covered on average only 11% of the detected rain duration (Figure 1a). However, the amount of rain in the convective intervals represented approximately 40% of the total detected rain (Figure 1b). This is due to the large difference in the rain intensities between the stratiform and convective rain (medians of 0.69 $mm \cdot h^{-1}$ (sd: 1.1 $mm \cdot h^{-1}$) and 4.49 $mm \cdot h^{-1}$ (sd: 8.5 $mm \cdot h^{-1}$), respectively) (Figure 1c).

In general, the proportions of the convective rain duration and the rain amount for convective and stratiform rain were similar for the 16 disdrometers, despite the differences in the accumulated rain durations and amounts detected by the 16 devices (as explained in Section 2.1).

However, the convective rain was not distributed evenly across GWLs; variations in the proportion of convective rain were detected both in the rain duration and the rain amount (Figure 2b,c). While some GWLs recorded higher than the average (11%) proportion of convective rain duration (e.g., HM, NZ, TRM, and TM), less than the average convective rain duration was measured in others (e.g., WS, NEZ, SEZ, and TRW) (Figure 2b). The variation between GWLs was higher regarding the rain amount (Figure 2c), with convective rain representing more than 65% of the rain amount for HM and NZ but 17% of WS and less than 2% for SEZ and HFZ.

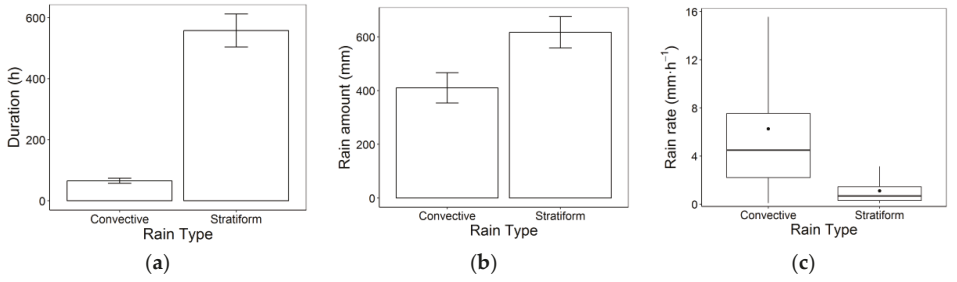


Figure 1. (**a**) A bar plot of accumulated rain duration in hours; (**b**) a bar plot of accumulated rain amount in mm; and (**c**) a boxplot of rain intensities in mm·h^{-1} for each rain type at Lausanne, Switzerland (December 2008–September 2010). Error bars in panels (**a**) and (**b**) are based on the standard error across devices. In panel (**c**) outliers are not shown, medians (means) are represented by the horizontal thick lines (black points), and the lower and upper hinges represent the first and third quartiles, while the upper (lower) whisker extends from the upper (lower) hinge to the largest (smallest) value no further than 1.5 times the interquartile range from the hinge [69].

Visual inspection of the convective rain proportions did not reveal differences between cyclonic and anticyclonic GWLs in duration or rain amount. Conversely, the flow direction revealed small convective rain duration proportions in the NE–E (6%) compared to higher percentages in the other flow directions, especially CE (17%) (Figure 3a). NE–E circulations were also associated with the smallest proportion of convective rain amount (31%) compared to the four other flow directions (37–51%) (Figure 3b). This was confirmed by comparing individual GWLs; the highest proportion of convective rain duration and amount occurred in HM, which belongs to GWLs with no specific flow direction CE, while no convective rain was recorded in some NE–E circulations such as HFZ and NEA (Figure 2b,c).

3.2. Rain Microstructure

3.2.1. Drop Size Distribution (DSD)

GWL-specific rain DSDs differed from each other and from the overall mean rain DSD. This is illustrated for both rain types by comparing the mean rain concentrations for each 1 mm of the diameter range (Figure 4). To aid in the interpretation, the plot in Figure 4 contains a selection of GWLs that demonstrated a distinct DSD shape.

In the case of stratiform rain (Figure 4a), most GWL-averaged concentrations followed the overall average concentration shape and fell within ±400 drops around the average rain concentration per diameter range. The variance was larger near the peak concentration, which corresponds to a drop size of 0.69 mm, and became smaller as the rain drop diameter diverged from the peak. The highest peaks of the stratiform rain concentration were associated with the GWLs SWA, NZ, and TM. SEZ had the lowest peak. The concentration curve had a broader peak shifted to the left side (a high number of small drops) for HNFZ, while HFA was slightly shifted to the right and recorded the highest number of drops larger than 1.5 mm in diameter. These differences led to significant differences between GWLs when assessing the rain microstructure parameters (see Section 3.2.2).

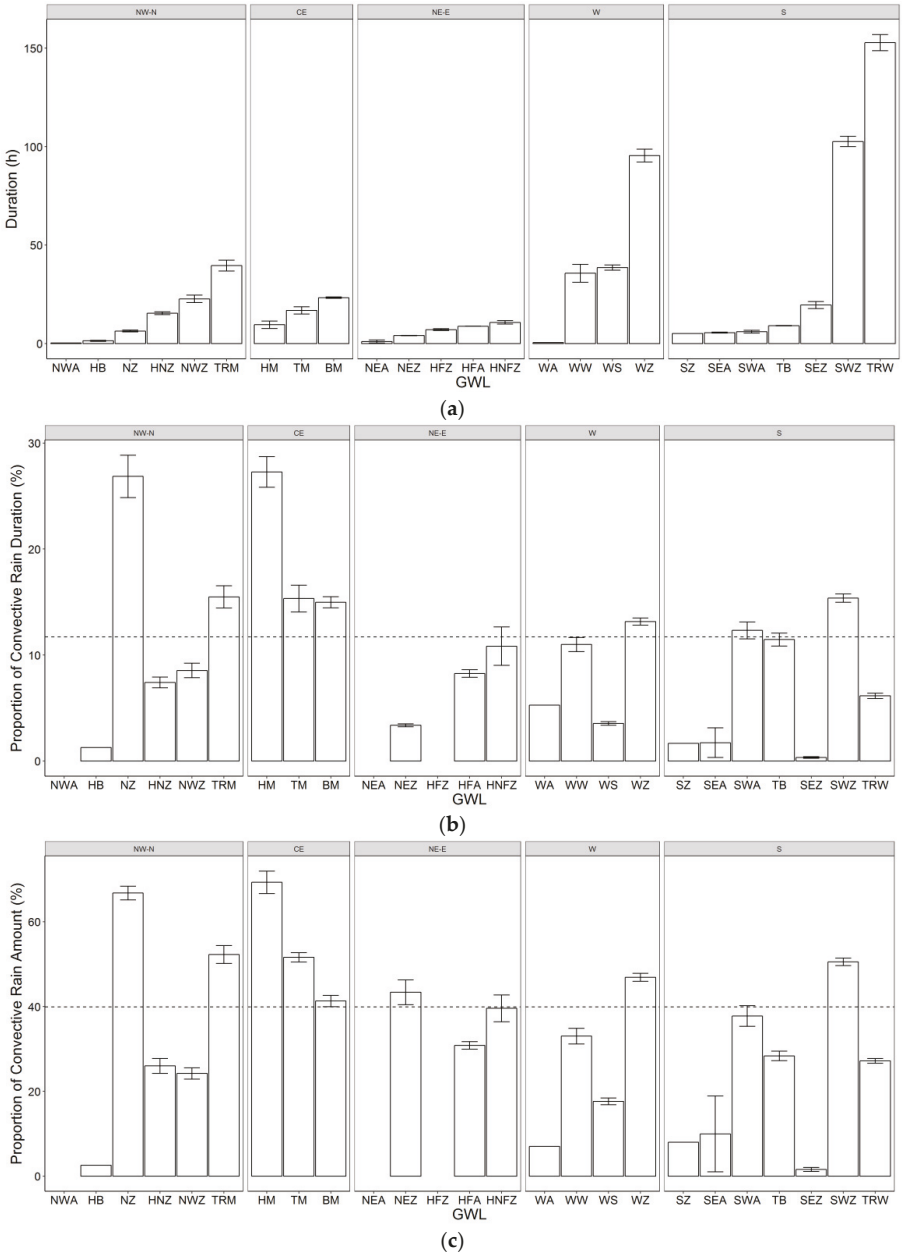


Figure 2. (**a**) Bar plots of the accumulated rain duration in hours; (**b**) proportion of the convective rain duration averaged over the 16 devices for each GWL (see Table 1); and (**c**) proportion of the convective rain amount averaged over the 16 devices for each GWL. Error bars are based on the standard error across devices. GWLs are grouped by flow direction and within each flow direction ordered by accumulated rain duration over all disdrometers. Dashed lines represent the overall proportion of the convective rain duration in panel (**b**) and the convective rain amount in panel (**c**).

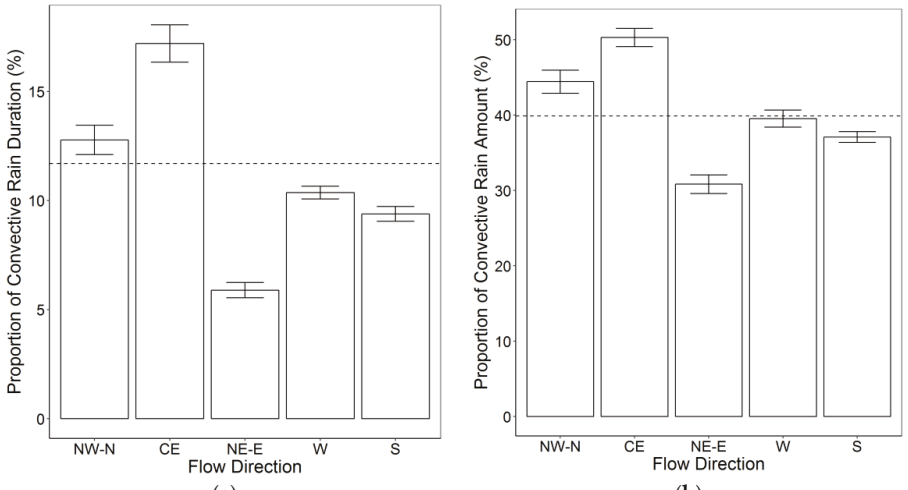


Figure 3. Proportion of convective rain (**a**) duration and (**b**) amount averaged over the 16 devices within each flow direction (cf. Table 1). Error bars are based on the standard error. Dashed lines represent the overall proportion of the convective rain duration and amount.

In contrast to the pronounced peak in stratiform rain DSD, convective rain showed a wider range of high concentrations (0.69–0.94 mm) (Figure 4b). The curves for convective rain were shifted toward larger diameters with much higher rain drop concentrations. The variation in the concentration averages between GWLs was much larger in all diameter ranges reaching up to ±1000 drops near the peak. The highest peaks were associated with the GWLs WS, SWA, TM, and NZ, with TM having the highest concentration of drops smaller than 1.5 mm in diameter. GWLs with the lowest concentration peaks were HFA, BM, TB, and WW. WW, HM, and HFA had the highest number of drops larger than 1.5 mm in diameter.

Testud et al. [70] proposed the concept of normalized distribution to compare rain DSDs without presumptions about their shapes. The results of applying the method of normalized DSD are provided in Figures S1 and S2.

3.2.2. Rain Parameters R, D_m and N

While the differences between GWLs regarding the rain DSD appear to be small, the high number of drops per interval and the large number of intervals led to significant differences in the three main rain parameters. The following comparisons were made without discrimination between the different devices.

When analyzing the stratiform rain intervals, the average values of R, N, and D_m of the three most frequent GWLs (TRW, SWZ, and WZ) were very close to the overall averages of these parameters (Figure 5a). These three GWLs are typical Atlantic frontal systems associated with moderate temperatures and therefore average rain intensities. GWLs with high average R (HFA, BM, WW, TB, and HM) recorded the highest D_m values, while GWLs with the highest N averages were associated with close to average R values. Conversely, GWLs with the smallest R averages (SEZ, NEZ, and HFZ) had low values of both average D_m and N. GWLs that belong to the same flow direction (identical colors in Figure 5a) are clustered in the space of N and D_m, with the exception of some outliers (HFA, SEZ, and SWA).

The R, D_m, and N values of stratiform rain vary significantly between the five classes of flow direction (cf. Table 1) despite their considerable variance within each flow direction. NE–E GWLs

had the smallest average value of R associated with the lowest average D_m (Figure 6). CE had the highest value of R with the highest value of N and a very high value of D_m. Both S and W had very high values of R and D_m.

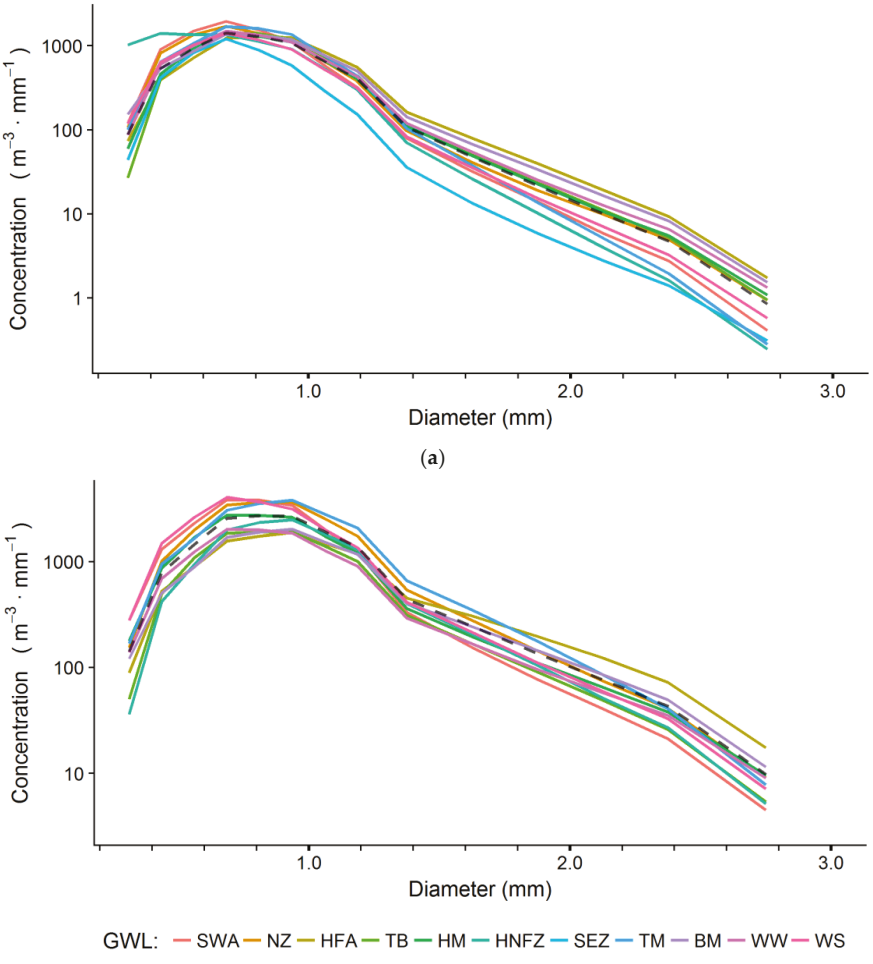


Figure 4. Mean concentrations (number of drops per 1 m^3 of air volume and per 1 mm of the diameter range (A4)) in (**a**) stratiform rain and (**b**) convective rain. To aid the interpretation, the plot is limited to a selection of GWLs that demonstrated distinct DSDs. The dashed black line indicates the mean DSD in both rain types regardless of GWL.

For convective rain, the three parameters R, D_m, and N had larger values and variations among GWLs than those for stratiform rain (Figure 5b). The average R values ranged from 3.67 mm·h^{-1} for NWZ to 8.56 mm·h^{-1} for HFA. High average R values were associated with large drops (D_m) and vice versa, with the exception of two cases: (1) TM had the second highest average R despite having a very low D_m, while N was comparably large and (2) WW had a less than average R value despite having the highest average D_m due to a small average N. The most frequent GWLs in convective rain (SWZ, WZ, and TRW) had higher than average R values but close to average D_m values.

The R, D_m and N values also varied for different flow directions in convective rain (Figure 6b). NW–N GWLs had the smallest rain intensity despite having the largest average N values. The other four flow directions recorded higher values of both D_m and R. This pattern differs from the one demonstrated in stratiform rain, which might be an indication of a larger influence of local conditions in convective rain.

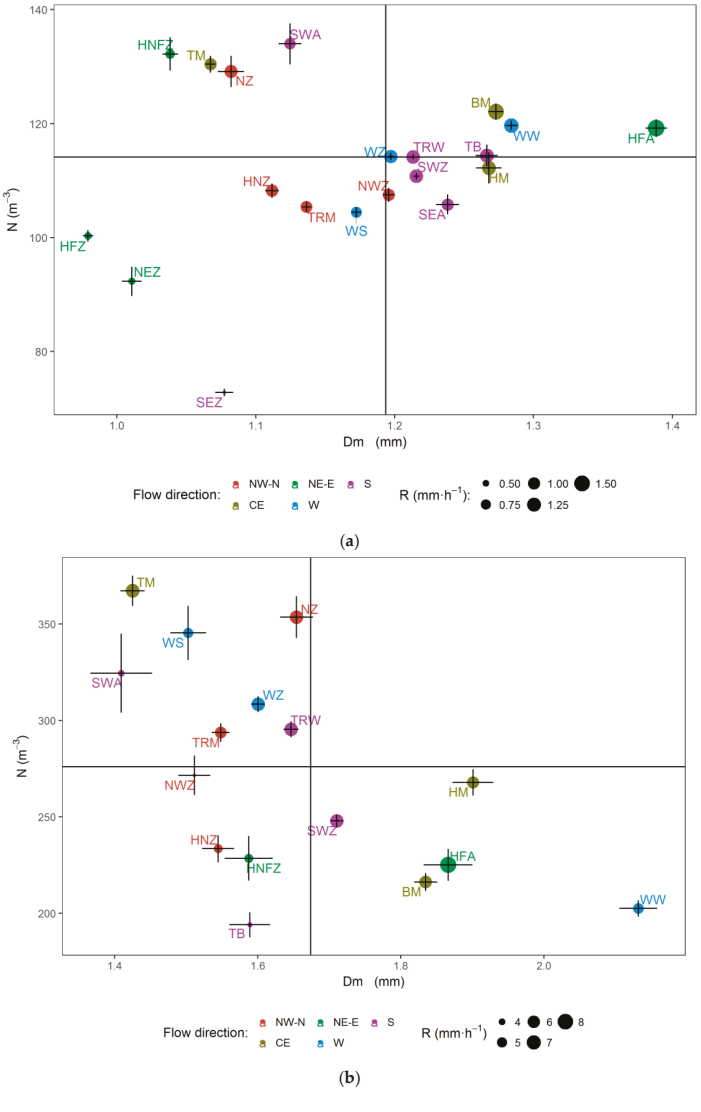


Figure 5. Scatter plot of the rain parameters per GWL with the mean mass weighted diameter (D_m) on the horizontal axis and the mean concentration (N) on the vertical axis. The size of each dot corresponds to the rain intensity (R) associated with a particular GWL. The colors indicate the five considered flow directions (cf. Table 1) in (**a**) stratiform rain and (**b**) convective rain. The horizontal lines represent the standard error of D_m. The vertical lines represent the standard error of N.

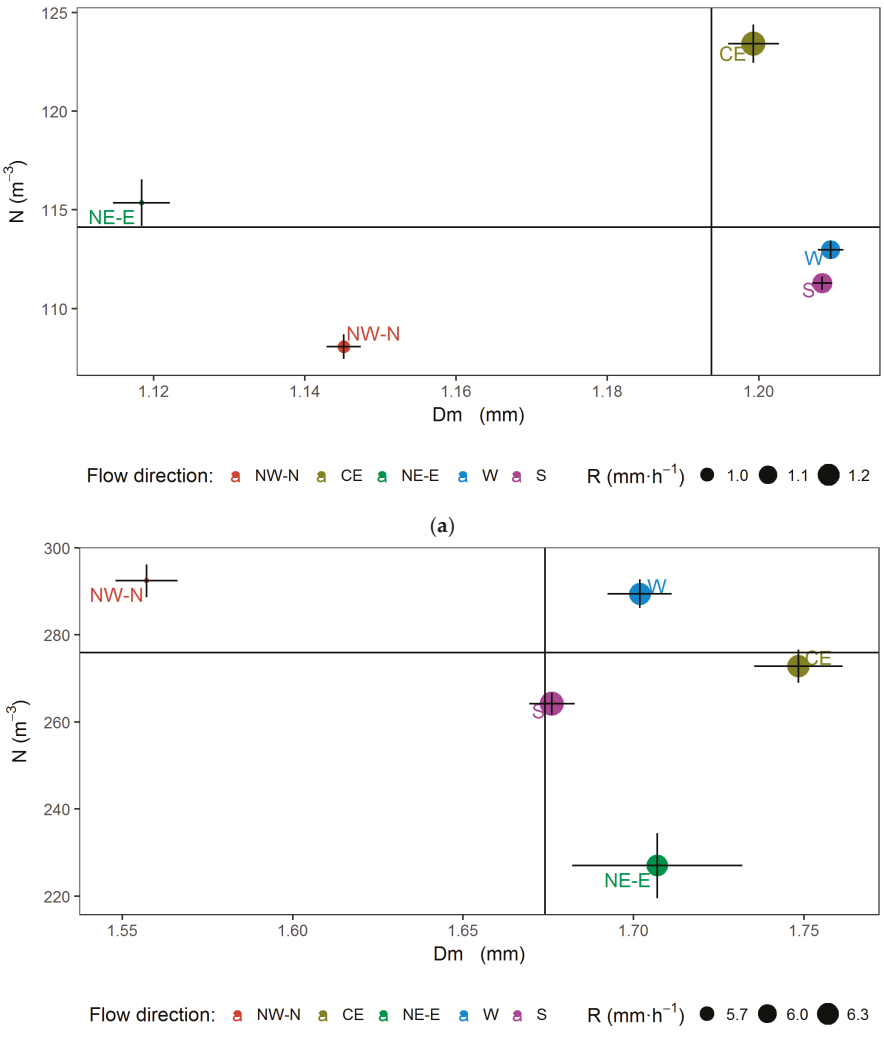


Figure 6. Scatter plot of rain parameters for each flow direction with the mean mass weighted diameter (D_m) for each flow direction on the horizontal axis and the mean concentration (N) for each flow direction on the vertical axis. The size of each dot represents the rain intensity associated with the corresponding flow direction. The colors indicate the five flow directions (see Table 1) in (**a**) stratiform rain and (**b**) convective rain. The horizontal lines represent the standard error of D_m. The vertical lines represent the standard error of N. Note the difference in absolute values for the two panels.

3.3. The Z–R Relation

Based on the global Z–R model for the entire dataset, A was estimated to be 295 for stratiform rain and 278 for convective rain. The values for b were 1.49 for stratiform rain and 1.54 for convective rain. For the GWL Z–R model, the A and b values differed between the individual GWLs in both rain types (Figure 7). Most A values for stratiform rain were clustered in the range of 215–350, while b values

clustered in the range of 1.35–1.65, with the exception of SEZ, which had higher values for both A and b. Convective rain had a larger range for both A (189–371) and b (1.36–1.84) compared to stratiform rain (excluding the SEZ results).

The spatial variation of the A and b values from a selection of events in this dataset has been discussed in detail by Jaffrain and Berne [71]. In our case, comparable results were achieved by using the entire dataset to estimate the A and b values associated with each of the 16 devices. The variability of the A and b values was significantly larger for the GWL Z–R model compared to the spatial variability in both rain types (Figure 7).

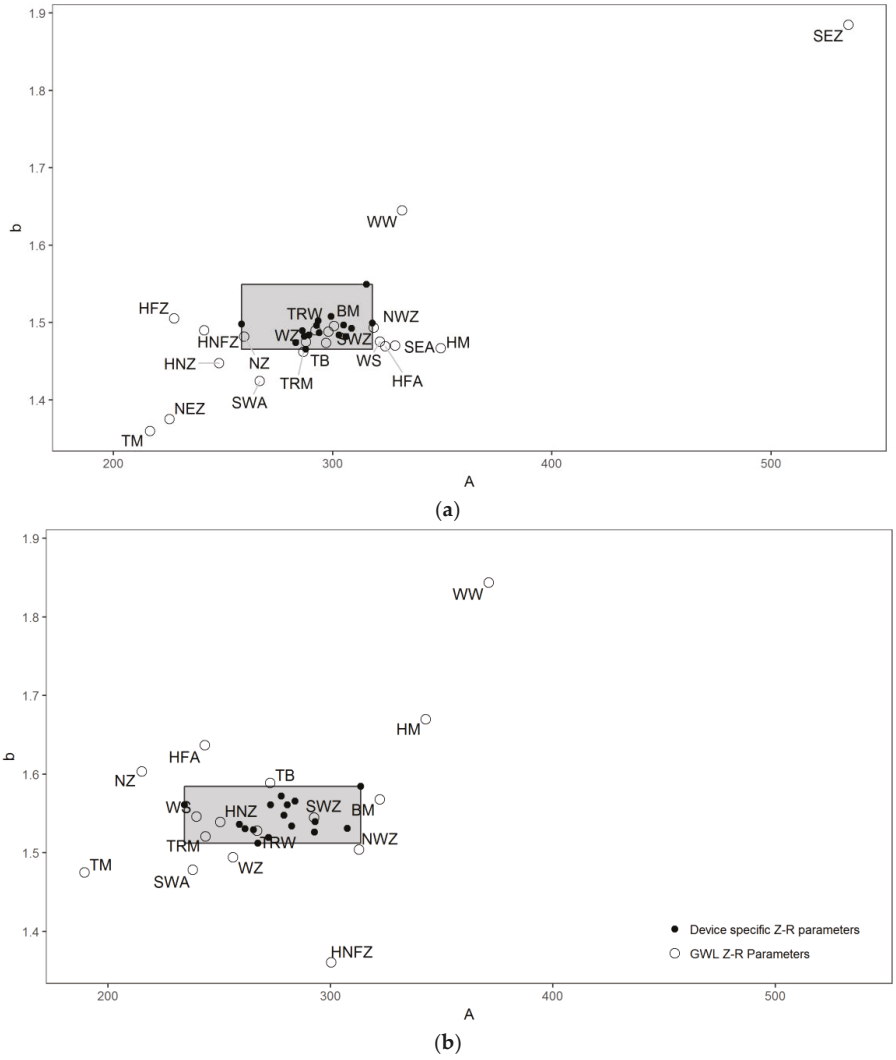


Figure 7. A and b values for the GWL Z–R model (white circles) and the 16 devices (dots) in (**a**) stratiform and (**b**) convective rain. The grey area represents the range of A and b values as they vary spatially (across the devices).

The mean average error (MAE) in the rain rates indicated a better performance of the GWL model in comparison to the global model (Table 2).

Table 2. Mean average error (MAE) in the rain rate R [mm·h^{-1}] calculated by the GWL Z–R model and the general Z–R relation. The values are given for the entire dataset and for each rain type separately.

	Whole Dataset	Stratiform Rain	Convective Rain
Global model	0.584	0.357	2.52
GWL model	0.55	0.345	2.3

Using the global model resulted in a 10.4% overestimation of the total detected rain amount, while using the GWL model resulted in an overestimation of 9.4%. The improvement resulting from using the GWL Z–R model was particularly high within some GWLs. For stratiform rain, the improvement reached 55%, 21%, and 11% for SEZ, WW, and HM, respectively. For convective rain, the improvement reached 92%, 47%, and 15% for WW, HM, and BM, respectively (Figure 8). These GWLs showed distinct averages of D_m and N, as illustrated in Figure 5.

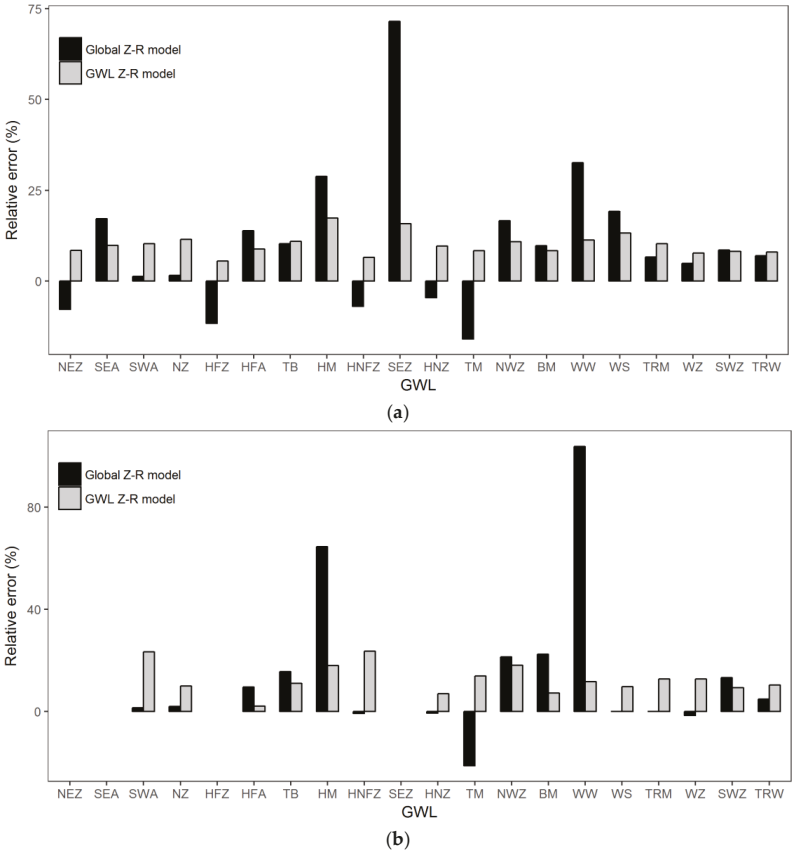


Figure 8. Relative error (RE) of the accumulated rain amount estimated by the global Z–R model (black) and the GWL Z–R model (grey) for (**a**) stratiform rain and (**b**) convective rain.

4. Discussion

4.1. Weather Type Classification

The Hess–Brezowsky classification was chosen due to its availability even for a longer period, its proven suitability for Central Europe, and its basis on the surface and 500 hPa height pressure patterns with the latter being especially important for the formation of precipitation [47]. The Hess–Brezowsky circulation type records experienced a shift in frequencies in the mid-1980s and therefore should not be used for climatological analyses [72]. However, this is not a problem in our case because the dataset and the analyses were limited to a period shorter than two years, which does not fall within the mid-1980s.

Another concern is the influence of orographic rain on the proportion of convective rain per GWL and per flow direction. The chosen method is not able to distinguish orographic rain. However, we assume that the proportion of orographic rain within a specific GWL or flow direction would be stable for a given location. Seasonality also might influence the proportion of convective rain per GWL. Unfortunately, the dataset used does not allow for a proper analysis of likely seasonal variation.

4.2. Rain Type Classification

Separation of convective and stratiform rain by the described simple method of [15] is prone to misclassification. This method does not recognize intervals where both convective and stratiform rain occur [73,74]. In addition, the transition between convective and stratiform rain cannot be properly distinguished by using the five successive 2-min rain intensity records [74]. However, it was chosen because data necessary for other separation methods were not available (e.g., wind profilers or 2D video disdrometers [75,76]).

The proportion of the convective rain amount in this dataset (40%) falls within the range reported for Central Europe by Rulfová and Kyselý [76] (30–50%). A higher percentage of convective rain is reported for tropical areas: on average 52% and 45% for continental and oceanic convective rain, respectively [77]. Even higher levels (52–69%) are reached for different locations in India [13].

Stratiform rain was more frequent than convective rain in all GWLs. However, convective rain represented more than half the total rain amount in some GWL classes, while this percentage dropped to less than 2% for other classes. GWLs with no specific prevailing flow direction (CE) provided a higher probability of convective rain due to their low-pressure gradients. Most GWLs associated with clear flow directions had a smaller proportion of convective rain because frontal structures may prevail. Despite the limited measurement duration of some of the addressed weather patterns, our results indicate that some GWLs provide conditions that favor one rain type more than the other.

Convective rain is expected to be more frequent with larger rain amounts in the future, especially because it is more sensitive than stratiform rain to the increase in temperature [78]. For example, convective rain has shown a positive trend in all seasons in the Czech Republic [79] and is expected to increase under climate change scenarios [80]. For Switzerland, there is an increase in the annual occurrence of convective weather types [81] and summer convective rain at high elevations [82]. The proportion of convective rain at elevated altitudes is expected to be higher in the future [83]. The reported trends in the proportion of convective rain, along with our results regarding the strong variation in the proportion of convective rain with GWL, highlight the importance of inspecting rain properties for both rain types separately. In other words, a comparison between general weather types without separating convective and stratiform rain may overestimate or underestimate the differences in the addressed rain properties. This error would be influenced by the different proportions of convective rain among these weather types.

4.3. Rain Microstructure

Differences in the rain microstructure were detected between GWL classes for both rain types. These differences were demonstrated by comparing the mean rain DSD and the parameters R, D_m, and N. While the reasons behind these variations for individual GWLs remain largely unclear, the flow directions

provide a partial explanation via the associated temperatures and the availability of condensation nuclei and water content.

Particularly in stratiform rain, western and southern circulations bring high water content from the Atlantic and the Mediterranean, which produces high rain intensity. The moisture flux in this area, and generally in the northern hemisphere above 30° N, is westerly and south westerly [84]. Aerosols influence cloud properties by acting as cloud condensation nuclei (CCN) [85]. The aerosol availability appears to be limited for the western, southern, and northwestern circulations in comparison to the northeastern circulations [86]. This is due to anthropogenic activities; a higher concentration of aerosols is reported when the sources are in Russia, Central Europe, and Eastern Europe [87]. Heavily industrialized areas in Germany, the Czech Republic, and Poland are also sources of high particle number concentrations [88]. Consequently, for the study area, the high water content and low CCN in western and southern circulation forces water particles to condensate on a smaller number of nuclei and makes rain drops comparatively large. Conversely, the northern flow direction is associated with low temperatures and forces air parcels to travel a longer distance over terrestrial areas; this lowers the average water content in comparison to other flow directions. Therefore, lower average values of R and D_m were observed. Only in the case of CE, where no flow direction is dominant, did R, D_m, and N exhibit high values, which might indicate a high availability of both CCN and water content. May et al. [89] reported a large number of small drops for high aerosol regimes, but an inverse relation between rain concentration and aerosol concentration in convective clouds. According to Cecchini et al. [90], an increase in CCN leads to higher rain drop concentrations in stratiform rain, and larger rain droplets in convective rain. This falls in line with our conclusions regarding NE–E GWLs, which have higher than average rain concentration in stratiform rain and larger than average rain drop size in convective rain.

Unfortunately, this does not explain the variations between individual GWLs that share the same flow direction. Grouping GWLs by flow direction is not a very precise procedure; the wind direction at higher levels might fluctuate to some extent even within individual GWLs when inspecting a specific location. This means that the source regions and back trajectories might be different within the same GWL [91]. In addition, as the GWL classes change, the site may be influenced by the previous or the upcoming GWL because the classification is done over a large scale. Furthermore, seasonal influences might play a role and there might be an effect of the main alpine ridge, especially for the south and east flow directions, as well as of the mountainous borders between Switzerland and France for the north and west flow directions [92–94].

Variations in the rain microstructure between GWLs and flow directions were also detected for convective rain. However, the patterns of these differences were different than the ones for stratiform rain. It might be that, during convective rain, the water content is already sufficiently high. R, N and D_m might be influenced largely by rain formation microscale processes, the height of freezing levels or the boundary layer relative humidity [24], and the local aspects such as topography [92]. All of these might act in combination with the influence of the general weather conditions and produce the observed variation. In other words, GWLs seem to provide conditions that stimulate specific regional and local processes, which enhance or suppress specific rain-forming processes, which in turn influence the observed rain DSD for both rain types.

In a study carried out in Leon, Spain, Fernandez-Raga et al. [53] found that rain drops associated with the northern flow direction had a smaller size compared to other flow directions. This is because most of the moisture transported from the north would precipitate before reaching the measurement site. Conversely, the high Atlantic moisture causes large rain drop size to be associated with western flows. Our results also reveal that GWLs associated with northern flow directions (NE–E and NW–N) had smaller rain drops compared to other flow directions, with the exception of NE–E for convective rain. Our results are in agreement despite using different classifications of the large-scale weather circulations and different locations within Europe. On the other hand, it appears in the reported results of Hachani et al. [29] that northeast circulations are characterized by low concentrations and large drops for the Cévennes–Vivarais region in France. The results also discuss the weak influence of the distance from the Mediterranean Sea,

the seasonal influence, the rain type influence, and a combination of these effects. However, the combined effect of rain type and synoptic weather patterns was not addressed, which might be the main reason for the different observation regarding northeast flow direction.

4.4. The Z–R Relation

Jaffrain and Berne [71] demonstrated the spatial variation of the A and b values for a selection of 36 events in the same dataset. In this study, we obtained similar values using the entire dataset. In addition, we demonstrated that the GWL-specific variation exceeds the 1-km^2 spatial variation of the A and b values for both rain types. This led to an improved estimation of the rain amount when accounting for the GWL variation over the entire measurement duration. This improvement was particularly high for some individual GWLs, especially those with distinct values of D_m and N, such as SEZ and WW in stratiform rain and WW and HM in convective rain. Interestingly, these GWLs were frequently encountered (Table 1), which rules out the concern of small sample size effect. The rain amount overestimation within 1 km^2, which can reach up to +15% as reported by Jaffrain and Berne [71], appears to be propagated to the rain amount estimation within individual GWLs.

The different synoptic features of GWLs at different locations (e.g., one station gets southerly maritime wind for one GWL whereas another gets eastward continental winds for the same GWL) may result in differing rain microstructures at these locations. Due to this local effect of GWLs, GWL models have to be calibrated for each station specifically.

Unfortunately, the number of intervals for each device within each GWL is not large enough to draw decisive conclusions about the spatial variation of the rain estimation within each GWL. A longer measurement period covering a larger number of rain events within each GWL may lead to more concrete conclusions, especially if associated with more precise wind directions at high elevations. Ground measurements over a wider area might help our understanding of the complexity of the interaction between the flow directions at different elevations and the topography and their influence on the rain microstructure.

5. Conclusions

Ground measurements of the rain drop size distribution (DSD) recorded by 16 disdrometers in the period of December 2008–September 2010 during the EPFL-LTE campaign were separated into convective and stratiform intervals and then combined with records of general weather patterns (GWL) to investigate the variations in the rain microstructure between the GWLs.

Convective rain proportions varied among GWLs in both duration and rain amount. These variations along with the previously established influence of rain type on rain DSD suggest that it is important to take the rain type into account when comparing GWLs.

The highest rain amounts and longest rain durations were recorded for the GWLs TRW, SWZ, and WZ. Consequently, these GWLs controlled the total averaged rain parameters. Nevertheless, other GWLs had different rain DSDs. The variation was stronger near the peak of the rain DSD. Significant differences were detected in the values of the rain intensity R, the mass weighted mean diameter D_m, and the rain concentration N. GWLs associated with high averaged R also recorded large rain drops and vice versa.

The values of R, D_m, and N were significantly different between flow directions. In the case of stratiform rain, northern flow (NE–E and NW–N) recorded the smallest average drop size and lowest average R values. S, W, and CE circulations had large drops and high R values. High numbers of drops were recorded in GWLs associated with CE and NE–E. Variations also exist in the case of convective rain; however, the pattern is different and might be affected by local conditions.

The parameters of the Z–R relation varied significantly between GWLs, and estimating R based on the GWL-specific Z–R relations improved the total rain amount estimation for the entire dataset and for both rain types. The improvement is highest for individual GWLs, which revealed distinct rain microstructures such as SEZ and WW in stratiform rain and WW and HM in convective rain.

To our knowledge, this is the first study in Central Europe linking rain DSD to synoptic weather types and general flow directions for each individual rain type. The effect of the described variations with GWL and the flow direction, e.g., on the Z–R parameters, might improve the current radar estimation of R. In future studies, the microstructure of extreme and heavy rain within each GWL, the role of the wind direction at high levels, and improved separation methods of convective and stratiform rain should be addressed to reveal more about the properties of rain within large-scale circulations. However, addressing the inherent spatial and temporal variability of the rain microstructure within each GWL calls for a larger dataset to provide sufficient rain measurements within each GWL and to support these results.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/10/6/811/s1, Table S1: "Accumulated rain duration and rain amount per disdrometer", Figure S1: "Normalized rain DSD for a selection of GWLs in stratiform rain", Figure S2: "Normalized rain DSD for a selection of GWLs in Convective rain".

Author Contributions: Wael Ghada and Annette Menzel conceptualized the research idea. Wael Ghada performed the data processing with support from Allan Buras and Marvin Lüpke. Wael Ghada wrote the manuscript. All authors contributed to the interpretation of results and the editing of the manuscript.

Acknowledgments: We thank those behind the EPFL-LTE campaign 2008–2010 for the published raw DSD records, the Deutscher Wetterdienst (German Meteorological Service-DWD) for providing records of GWLs, and Clemens Wastl for his insight regarding the influence of GWLs. This work was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program. The first author thanks the Deutscher Akademischer Austauschdienst (DAAD) for financial support. We appreciate the valuable comments provided by four anonymous reviewers.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

Appendix A

The following equations were used to calculate the key rain properties for each time step based on the disdrometer raw output.

1. Rain intensity (R) ($mm \cdot h^{-1}$) [16]

$$R = 60 \times 60 \times \pi / (6 \times 1{,}000{,}000\ \Delta T) \sum_{i=1}^{i=32} \sum_{j=1}^{j=32} \left(x_{i,j}\ {D_i}^3 / A_i \right) \quad \text{(A1)}$$

$x_{i,j}$: 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 (m^2): Corrected detection area [60]: $A_i = 180 \times (30 - D_i/2)/1{,}000{,}000$.

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

2. Reflectivity (Z) ($mm^6 \cdot m^{-3}$) [16]

$$Z = \sum_{i=1}^{i=32} \sum_{j=1}^{j=32} \left(x_{i,j}\ {D_i}^6 / (A_i\ V_j \Delta T) \right) \quad \text{(A2)}$$

$V_j \left(m \cdot s^{-1} \right)$: Mean velocity of drops in velocity range j.

The reflectivity is usually expressed in another unit [dBZ]:

$$Z_{[dBZ]} = 10 \times \log_{10} \left(Z_{[mm^6 \cdot\ m^{-3}]} \right). \quad \text{(A3)}$$

3. Total drop concentration (N) (m^{-3}), i.e., the total number of drops per m^3 [16]:

$$N = \sum_{i=1}^{i=32} \sum_{j=1}^{j=32} x_{i,j} / \left(V_j \, \Delta T \, W_i \, A_i \right) \tag{A4}$$

W_i (mm): Width of a diameter range i.

4. Mass weighted diameter (D_m) (mm) [61]:

$$D_m = \frac{\sum_{i=1}^{i=32} \sum_{j=1}^{j=32} \left(x_{i,j} \, D_i{}^4 \right)}{\sum_{i=1}^{i=32} \sum_{j=1}^{j=32} \left(x_{i,j} \, D_i{}^3 \right)}. \tag{A5}$$

References

1. 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]
2. Lal, R. *Soil Erosion Research Methods*, 2nd ed.; Soil and Water Conservation Society; St. Lucie Press: Ankeny, IA, USA, 1994.
3. 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]
4. Murphy, L.U.; Cochrane, T.A.; O'Sullivan, A. Build-up and wash-off dynamics of atmospherically derived Cu, Pb, Zn and TSS in stormwater runoff as a function of meteorological characteristics. *Sci. Total Environ.* **2015**, *508*, 206–213. [CrossRef] [PubMed]
5. Moupfouma, F.; Martin, L. Modelling of the rainfall rate cumulative distribution for the design of satellite and terrestrial communication systems. *Int. J. Satell. Commun. Netw.* **1995**, *13*, 105–115. [CrossRef]
6. Shkarofsky, I. Dependence of rain attenuation and cross-polarization on drop size distribution. *IEEE Trans. Antennas Propag.* **1979**, *27*, 538–542. [CrossRef]
7. 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]
8. 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]
9. Tapiador, F.J.; Navarro, A.; Moreno, R.; Jiménez-Alcázar, A.; Marcos, C.; Tokay, A.; Durán, L.; Bodoque, J.M.; Martín, R.; Petersen, W. On the Optimal Measuring Area for Pointwise Rainfall Estimation: A Dedicated Experiment with 14 Laser Disdrometers. *J. Hydrometeorol.* **2017**, *18*, 753–760. [CrossRef]
10. 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*, W00H06. [CrossRef]
11. Jaffrain, J.; Berne, A. Quantification of the Small-Scale Spatial Structure of the Raindrop Size Distribution from a Network of Disdrometers. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 941–953. [CrossRef]
12. 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]
13. Das, S.; Maitra, A. Characterization of tropical precipitation using drop size distribution and rain rate-radar reflectivity relation. *Theor. Appl. Climatol.* **2018**, *132*, 275–286. [CrossRef]
14. 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. Meteorol. Climatol.* **2015**, *54*, 1069–1089. [CrossRef]
15. 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]
16. 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. Meteorol. Climatol.* **2016**, *55*, 621–634. [CrossRef]

17. Maki, M.; Keenan, T.D.; Sasaki, Y.; Nakamura, K. Characteristics of the Raindrop Size Distribution in Tropical Continental Squall Lines Observed in Darwin, Australia. *J. Appl. Meteorol.* **2001**, *40*, 1393–1412. [CrossRef]
18. Islam, T.; Rico-Ramirez, M.A.; Thurai, M.; Han, D. Characteristics of raindrop spectra as normalized gamma distribution from a Joss–Waldvogel disdrometer. *Atmos. Res.* **2012**, *108*, 57–73. [CrossRef]
19. Hu, Z.; Srivastava, R.C. Evolution of Raindrop Size Distribution by Coalescence, Breakup, and Evaporation: Theory and Observations. *J. Atmos. Sci.* **1995**, *52*, 1761–1783. [CrossRef]
20. Villermaux, E.; Bossa, B. Single-drop fragmentation determines size distribution of raindrops. *Nat. Phys.* **2009**, *5*, 697–702. [CrossRef]
21. Kumjian, M.R.; Ryzhkov, A.V. The Impact of Evaporation on Polarimetric Characteristics of Rain: Theoretical Model and Practical Implications. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 1247–1267. [CrossRef]
22. Rosenfeld, D.; Mintz, Y. Evaporation of Rain Falling from Convective Clouds as Derived from Radar Measurements. *J. Appl. Meteorol.* **1988**, *27*, 209–215. [CrossRef]
23. Kumjian, M.R.; Ryzhkov, A.V. The Impact of Size Sorting on the Polarimetric Radar Variables. *J. Atmos. Sci.* **2012**, *69*, 2042–2060. [CrossRef]
24. 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]
25. Berg, W.; L'Ecuyer, T.; van den Heever, S. Evidence for the impact of aerosols on the onset and microphysical properties of rainfall from a combination of satellite observations and cloud-resolving model simulations. *J. Geophys. Res.* **2008**, *113*. [CrossRef]
26. Givati, A.; Rosenfeld, D. Separation between Cloud-Seeding and Air-Pollution Effects. *J. Appl. Meteorol.* **2005**, *44*, 1298–1314. [CrossRef]
27. Harikumar, R. Orographic effect on tropical rain physics in the Asian monsoon region. *Atmos. Sci. Lett.* **2016**, *17*, 556–563. [CrossRef]
28. 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]
29. Hachani, S.; Boudevillain, B.; Delrieu, G.; Bargaoui, Z. Drop Size Distribution Climatology in Cévennes-Vivarais Region, France. *Atmosphere* **2017**, *8*. [CrossRef]
30. Wen, J.; Zhao, K.; Huang, H.; Zhou, B.; Yang, Z.; Chen, G.; Wang, M.; Wen, L.; Dai, H.; Xu, L.; et al. Evolution of microphysical structure of a subtropical squall line observed by a polarimetric radar and a disdrometer during OPACC in Eastern China. *J. Geophys. Res.* **2017**, *122*, 8033–8050. [CrossRef]
31. Steiner, M.; Smith, J.A. Convective versus stratiform rainfall: An ice-microphysical and kinematic conceptual model. *Atmos. Res.* **1998**, *47-48*, 317–326. [CrossRef]
32. 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*; American Meteorological Society: Boston, MA, USA, 2003.
33. Campos, E.F.; Zawadzki, I.; Petitdidier, M.; Fernández, W. Measurement of raindrop size distributions in tropical rain at Costa Rica. *J. Hydrol.* **2006**, *328*, 98–109. [CrossRef]
34. Penide, G.; Kumar, V.V.; Protat, A.; May, P.T. Statistics of Drop Size Distribution Parameters and Rain Rates for Stratiform and Convective Precipitation during the North Australian Wet Season. *Mon. Weather Rev.* **2013**, *141*, 3222–3237. [CrossRef]
35. 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] [PubMed]
36. Ramos, A.M.; Barriopedro, D.; Dutra, E. Circulation weather types as a tool in atmospheric, climate, and environmental research. *Front. Environ. Sci.* **2015**, *3*. [CrossRef]
37. Brdossy, A.; Caspary, H.J. Detection of climate change in Europe by analyzing European atmospheric circulation patterns from 1881 to 1989. *Theor. Appl. Climatol.* **1990**, *42*, 155–167. [CrossRef]
38. Baur, F.; Hess, P.; Nagel, H. Kalender der Grosswetterlagen Europas 1881–1939. *Bad Homburg* **1944**, 35.
39. Gerstengarbe, F.-W.; Werner, P.C.; Busold, W.; Hess, P.; Brezowsky, H. *Katalog der Grosswetterlagen Europas nach Paul Hess und Helmuth Brezowski 1881–1992*; 4., vollständig neu bearbeitete Aufl.; Selbstverlag des Deutschen Wetterdienstes: Offenbach am Main, Germany, 1993; p. 244. Available online: https://www.dwd.de/DE/leistungen/pbfb_verlag_berichte/pdf_einzelbaende/113_pdf.pdf?__blob=publicationFile&v=3 (accessed on 23 May 2018).

40. BUISHAND, T.A.; BRANDSMA, T. Comparison of circulation classification schemes for predicting temperature and precipitation in the Netherlands. *Int. J. Climatol.* **1997**, *17*, 875–889. [CrossRef]
41. Cony, M.; Martín, L.; Hernándes, E.; Del Teso, T. 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 9 April 2018).
42. 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]
43. Jacobeit, J.; Glaser, R.; Luterbacher, J.; Wanner, H. Links between flood events in central Europe since AD 1500 and large-scale atmospheric circulation modes. *Geophys. Res. Lett.* **2003**, *30*. [CrossRef]
44. 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]
45. 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]
46. 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]
47. Ustrnul, Z.; Czekierda, D. Circulation background of the atmospheric precipitation in Central Europe (based on the Polish example). *Meteorol. Z.* **2001**, *10*, 103–111. [CrossRef]
48. Ducić, V.; Luković, J.; Burić, D.; Stanojević, G.; Mustafić, S. Precipitation extremes in the wettest Mediterranean region (Krivošije) and associated atmospheric circulation types. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 687–697. [CrossRef]
49. Nowosad, J.; Stach, A. Relation between extensive extreme precipitation in Poland and atmospheric circulation. *Quaest. Geogr.* **2014**, *33*, 115–129. [CrossRef]
50. 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]
51. Hoy, A.; Schucknecht, A.; Sepp, M.; Matschullat, J. Large-scale synoptic types and their impact on European precipitation. *Theor. Appl. Climatol.* **2014**, *116*, 19–35. [CrossRef]
52. Blenkinsop, S.; Chan, S.C.; Kendon, E.J.; Roberts, N.M.; Fowler, H.J. Temperature influences on intense UK hourly precipitation and dependency on large-scale circulation. *Environ. Res. Lett.* **2015**, *10*. [CrossRef]
53. Fernandez-Raga, M.; Castro, A.; Marcos, E.; Palencia, C.; Fraile, R. Weather types and rainfall microstructure in Leon, Spain. *Int. J. Climatol.* **2017**, *37*, 1834–1842. [CrossRef]
54. Berne, A.; Jaffrain, J. *GPM Ground Validation NASA EPFL-LTE Parsivel DSD Data Lausanne, Switzerland*; NASA Global Hydrology Resource Center DAAC: Huntsville, AL, USA, 2016.
55. Knittel, J. Großwetterlage: Januar bis Dezember 2008. Available online: https://www.dwd.de/DE/leistungen/grosswetterlage/2008/gwl_zusammenfassung.pdf?__blob=publicationFile&v=3 (accessed on 3 April 2018).
56. Knittel, J. Großwetterlage: Januar bis Dezember 2009. Available online: https://www.dwd.de/DE/leistungen/grosswetterlage/2009/gwl_zusammenfassung.pdf?__blob=publicationFile&v=3 (accessed on 3 April 2018).
57. Knittel, J. Großwetterlage: Januar bis Dezember 2010. Available online: https://www.dwd.de/DE/leistungen/grosswetterlage/2010/gwl_zusammenfassung.pdf?__blob=publicationFile&v=3 (accessed on 3 April 2018).
58. OTT MESSTECHNIK GmbH & Co. K. Present Weather Sensor Parsivel: Operating Instructions. Available online: https://www.esrl.noaa.gov/psd/data/obs/instruments/OpticalDisdrometer.pdf (accessed on 29 March 2018).
59. Löffler-Mang, M.; Joss, J. An Optical Disdrometer for Measuring Size and Velocity of Hydrometeors. *J. Atmos. Ocean. Technol.* **2000**, *17*, 130–139. [CrossRef]
60. 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]
61. Marzuki, M.; Randeu, W.L.; Schönhuber, M.; Bringi, V.N.; Kozu, T.; Shimomai, T. Raindrop Size Distribution Parameters of Distrometer Data With Different Bin Sizes. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3075–3080. [CrossRef]
62. Marzano, F.S.; Cimini, D.; Montopoli, M. Investigating precipitation microphysics using ground-based microwave remote sensors and disdrometer data. *Atmos. Res.* **2010**, *97*, 583–600. [CrossRef]
63. Tokay, A.; Petersen, W.A.; Gatlin, P.; Wingo, M. Comparison of Raindrop Size Distribution Measurements by Collocated Disdrometers. *J. Atmos. Ocean. Technol.* **2013**, *30*, 1672–1690. [CrossRef]

64. Friedrich, K.; Kalina, E.A.; Masters, F.J.; Lopez, C.R. Drop-Size Distributions in Thunderstorms Measured by Optical Disdrometers during VORTEX2. *Mon. Weather Rev.* **2013**, *141*, 1182–1203. [CrossRef]
65. Atlas, D.; Srivastava, R.C.; Sekhon, R.S. Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.* **1973**, *11*, 1–35. [CrossRef]
66. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]
67. Team, R.C. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017; Available online: https://www.R-project.org/ (accessed on 22 May 2018).
68. Team, R. *RStudio. Integrated Development Environment for R*; RStudio Inc.: Boston, MA, USA, 2016; Available online: http://www.rstudio.com/ (accessed on 22 May 2018).
69. Wickham, H. *ggplot2. Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2009.
70. 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]
71. Jaffrain, J.; Berne, A. Influence of the Subgrid Variability of the Raindrop Size Distribution on Radar Rainfall Estimators. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 780–785. [CrossRef]
72. Cahynová, M.; Huth, R. Enhanced lifetime of atmospheric circulation types over Europe: Fact or fiction? *Tellus A* **2009**, *61*, 407–416. [CrossRef]
73. 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]
74. 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]
75. Williams, C.R.; Ecklund, W.L.; Gage, K.S. Classification of Precipitating Clouds in the Tropics Using 915-MHz Wind Profilers. *J. Atmos. Ocean. Technol.* **1995**, *12*, 996–1012. [CrossRef]
76. Rulfová, Z.; Kyselý, J. Trends of Convective and Stratiform Precipitation in the Czech Republic, 1982–2010. *Adv. Meteorol.* **2014**, *2014*, 1–11. [CrossRef]
77. Yang, S.; Smith, E.A. Convective–Stratiform Precipitation Variability at Seasonal Scale from 8 Yr of TRMM Observations: Implications for Multiple Modes of Diurnal Variability. *J. Clim.* **2008**, *21*, 4087–4114. [CrossRef]
78. Berg, P.; Moseley, C.; Haerter, J.O. Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci* **2013**, *6*, 181–185. [CrossRef]
79. Rulfová, Z.; Kyselý, J. Disaggregating convective and stratiform precipitation from station weather data. *Atmos. Res.* **2013**, *134*, 100–115. [CrossRef]
80. Rulfová, Z.; Beranová, R.; Kyselý, J. Climate change scenarios of convective and large-scale precipitation in the Czech Republic based on EURO-CORDEX data. *Int. J. Climatol.* **2017**, *37*, 2451–2465. [CrossRef]
81. Stefanicki, G.; Talkner, P.; Weber, R.O. Frequency Changes of Weather Types in the Alpine Region since 1945. *Theor. Appl. Climatol.* **1998**, *60*, 47–61. [CrossRef]
82. Giorgi, F.; Torma, C.; Coppola, E.; Ban, N.; Schär, C.; Somot, S. Enhanced summer convective rainfall at Alpine high elevations in response to climate warming. *Nat. Geosci* **2016**, *9*, 584–589. [CrossRef]
83. Fischer, A.M.; Keller, D.E.; Liniger, M.A.; Rajczak, J.; Schär, C.; Appenzeller, C. Projected changes in precipitation intensity and frequency in Switzerland: A multi-model perspective. *Int. J. Climatol.* **2015**, *35*, 3204–3219. [CrossRef]
84. 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*. [CrossRef]
85. Lohmann, U.; Feichter, J. Global indirect aerosol effects: A review. *Atmos. Chem. Phys.* **2005**, *5*, 715–737. [CrossRef]
86. Ackermann, I.J.; Hass, H.; Memmesheimer, M.; Ebel, A.; Binkowski, F.S.; Shankar, U. Modal aerosol dynamics model for Europe. *Atmos. Environ.* **1998**, *32*, 2981–2999. [CrossRef]
87. 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]
88. 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]
89. 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]

90. 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]
91. Hondula, D.M.; Sitka, L.; Davis, R.E.; Knight, D.B.; Gawtry, S.D.; Deaton, M.L.; Lee, T.R.; Normile, C.P.; Stenger, P.J. A back-trajectory and air mass climatology for the Northern Shenandoah Valley, USA. *Int. J. Climatol.* **2010**, *30*, 569–581. [CrossRef]
92. Fuhrer, O.; Schär, C. Embedded Cellular Convection in Moist Flow past Topography. *J. Atmos. Sci.* **2005**, *62*, 2810–2828. [CrossRef]
93. Cotton, W.R.; Bryan, G.; van den Heever, S.C. The Influence of Mountains on Airflow, Clouds, and Precipitation. In *Storm and Cloud Dynamics: The Dynamics of Clouds and Precipitating Mesoscale Systems*, 2nd ed.; Cotton, W.R., Bryan, G.H., van den Heever, S.C., Eds.; Academic Press: Burlington, MA, USA, 2011; pp. 673–750.
94. Rotunno, R.; Houze, R.A. Lessons on orographic precipitation from the Mesoscale Alpine Programme. *Q. J. R. Meteorol. Soc.* **2007**, *133*, 811–830. [CrossRef]

remote sensing

MDPI

Article

How Well Can IMERG Products Capture Typhoon Extreme Precipitation Events over Southern China?

Chaoying Huang [1], Junjun Hu [2,*], Sheng Chen [3,*], Asi Zhang [3], Zhenqing Liang [1], Xinhua Tong [1], Liusi Xiao [3], Chao Min [3] and Zengxin Zhang [4]

[1] Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Guangxi Teachers Education University, Ministry of Education, Nanning 530011, China; huang.chaoying@163.com (C.H.); liangzhenqing163@163.com (Z.L.); tongxh001@163.com (X.T.)

[2] Cooperative Institute for Mesoscale Meteorological Studies, The University of Oklahoma, Norman, OK 73072, USA

[3] School of Atmospheric Sciences, and Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Guangzhou 510275, China; zhangas@mail2.sysu.edu.cn (A.Z.); xiaoliusi104@163.com (L.X.); minchao1109@163.com (C.M.)

[4] State Key Laboratory of Hydrology-Water Resources and Hydraulics Engineering, Hohai University, Nanjing 210098, China; nfuzhang@163.com

* Correspondence: Junjun.Hu-1@ou.edu (J.H.); chensheng@mail.sysu.edu.cn (S.C.); Tel.: +1-405-219-4624 (J.H.); +86-0756-3668-330 (S.C.)

Received: 3 October 2018; Accepted: 21 December 2018; Published: 2 January 2019

Abstract: This study assesses the performance of the latest version 05B (V5B) Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (GPM) (IMERG) Early and Final Runs over southern China during six extremely heavy precipitation events brought by six powerful typhoons from 2016 to 2017. Observations from a dense network composed of 2449 rain gauges are used as reference to quantify the performance in terms of spatiotemporal variability, probability distribution of precipitation rates, contingency scores, and bias analysis. The results show that: (1) both IMERG with gauge calibration (IMERG_Cal) and without gauge correction (IMERG_Uncal) generally capture the spatial patterns of storm-accumulated precipitation with moderate to high correlation coefficients (CCs) of 0.57–0.87, and relative bias (RB) varying from −17.21% to 30.58%; (2) IMERG_Uncal and IMERG_Cal capture well the area-average hourly series of precipitation over rainfall centers with high CCs ranging from 0.78 to 0.94; (3) IMERG_Cal tends to underestimate precipitation especially the rainfall over the rainfall centers when compared to IMERG_Uncal. The IMERG Final Run shows promising potentials in typhoon-related extreme precipitation storm applications. This study is expected to give useful feedbacks about the latest V5B Final Run IMERG product to both algorithm developers and the scientific end users, providing a better understanding of how well the V5B IMERG products capture the typhoon extreme precipitation events over southern China.

Keywords: typhoon; IMERG; GSMaP; Southern China

1. Introduction

Extreme weather events are more and more frequent in recent decades [1–4]. In 2016, there were 12 typhoons that brought tremendous precipitation to China. In 2017, both China and United States were successively hit by a couple of tropical cyclones. In 15 days from 22 August to 5 September 2017, southern China were severely hit by three successive powerful typhoons, including the super typhoon Hato, the most violent cyclone since 1949 with maximum wind speed about 66.9 m/h above wind scale 17 (http://www.mca.gov.cn/article/yw/jzjz/zqkb/201708/20170800005628.shtml). The typhoon Hato not only brought tremendous rainfall to Guangdong province but also induced a severe storm

surge to the coastal region in Guangdong province. Five adjacent provinces (i.e., Guangdong, Guangxi, Fujian, Guizhou, and Yunnan) in southern China as well as Macao and Hong Kong were affected by typhoon Hato. The typhoon Hato led to 16 deaths and $6.82 billion in direct damages over China.

Tropical cyclone usually forms over tropical and sub-tropical oceans and grows gradually with diameters ranging from 100 km to 2000 km [5,6]. It is difficult to observe such a big weather system with traditional instruments like gauge and ground-based radar. With the development of remote sensing technology, satellite has become an important platform to provide high resolution continuous observations of precipitation at large regional scale or global scale. Satellite-based quantitative precipitation estimation (QPE) products emerged in recent decades. The popular satellite QPE products include Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [7] based on infrared (IR) sensor observations, PERSIANN-Cloud Classification System (PERSIANN-CCS) [8], the Climate Prediction Center morphing method (CMORPH) [9] based on passive microwave (PMW) sensor observations, the Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) [10,11] based on IR and PWM sensor measurements, and Global Satellite Mapping of Precipitation project (GSMaP) [12–15] derived from IR and PMW sensor observations. These products were built upon the success of Tropical Rainfall Measuring Mission (TRMM) launched on 28 November 1997, and have been widely used in various applications like hydrology, meteorology, climate change, and other precipitation-related researches [16–22].

After the retirement of TRMM in 15 April 2015, satellite-based QPE entered the Global Precipitation Measurement (GPM) era. GPM was initiated by U.S. National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA) with an aim to improve the accuracy of QPE and tocover a larger area than those of TRMM [23]. Upon the success of TRMM, a new QPE product, i.e., Integrated Multi-satellitE Retrievals for GPM (IMERG), was developed and released recently. IMERG takes the advantage of PERSIANN, CMORPH, and TMPA. It provides the public with a suite of QPE products with high resolution of 10 km/30 min, including Early (Late) run multi-satellite products ~4 (12) hours after observation time, and Final Run multi-satellite product ~2.5 months after the observation month once the monthly gauge analysis is received. The latest version of IMERG is 05B (V5B) released on 1 December 2017. Lots of reports have been found to investigate the performance of PERSIANN, CMORPH, TMPA, and GSMaP in the past decade [17,21,24–27]. Recently, a lot of studies have been reported in assessing the performance of the IMERG products in previous versions [28–36]. Most of previous studies focused on annual, seasonal, and diurnal characteristics of IMERG products over large regions like China and Continuous United States (CONUS), or small watershed basins. Recently, a few studies evaluated the performance of latest V5B IMERG products [37–41]. The study by Omranian et al. (2018) investigates the error characteristics of V5B IMERG Final Run product during the historical devastating hurricane Harvey [38]. Zhang et al. (2018) evaluated the performance of V5B IMERG final run products and Version 4 (V4) Global Satellite Mapping of Precipitation (GSMaP) products during a 60-year return extreme precipitation storm on 7 May 2017 over southern China [40]. However, to our best knowledge, few efforts have been reported in assessing the performance of the V5B IMERG during typhoon-related extreme precipitation events over China.

The objective of this study is to assess the performances of V5B IMERG Early Run and Final Run products during six typhoon-related extreme precipitation events over southern China in 2016 and 2017 (Figure 1). The detailed information of these events are given in Table 1. The assessment was carried out at hourly scale with spatial resolution of $0.1^\circ \times 0.1^\circ$ using graphical and statistical methods. The hourly observations from 2449 rain gauges were used as reference to quantify the performance in terms of spatiotemporal variability, probability distribution of precipitation rates, and contingency scores. In order to investigate the error source of IMERG, three types of bias (i.e., hit bias (HB), miss bias (MB), and false bias (FB)) are examined to quantify the contribution of different error sources, thus providing useful information for IMERG users and algorithm developers.

The rest of this paper is organized as follows: Section 2 describes the study region, tracks of six typhoons in this study, and available precipitation products to be evaluated. Section 3 focuses

on analyzing IMERG performances characterized by spatial distribution of precipitation, hourly precipitation time-series analysis, probability of detection of precipitation, agreement and disagreement analysis. A brief summary and conclusions are given in Section 4.

Table 1. Specific information of six typhoon events.

No.	Typhoon Name	Start Time (UTC)	End Time (UTC)	Total Duration (h)	Max Accumulated Gauged Rainfall (mm)	Max Gauged Rain Rate (mm/h)
1	Nida	2016-08-01 04:00	2016-08-03 06:00	50	333.8	106.9
2	Haima	2016-10-20 16:00	2016-10-21 23:00	32	244.5	76.5
3	Merbok	2017-06-12 00:00	2017-06-14 14:00	62	361	103.6
4	Hato	2017-08-22 00:00	2017-08-25 00:00	72	391.3	126.3
5	Pakhar	2017-08-26 00:00	2017-08-29 00:00	72	417.3	151.3
6	Mawar	2017-09-03 00:00	2017-09-06 00:00	72	284	97.6

2. Study Region, Data and Method

2.1. Study Region

The study region is Guangdong province located in southern China with longitude ranging from 109.66°E to 117.19°E and latitude varying between 20.30°N and 25.52°N (Figure 1a). The study region features with complex mountains and hills in the north, the northeast and the southwest, encompassing Pearl River Delta (PRD) indicated by red circle in Figure 1a. The flat PRD looks like a horn and is located in the central study region, which benefits the formation of precipitation in this area. Small flat plains scatter in coastal regions in the southwest and the northeast. Climatologically, this region is warm and humid, and suffers from massive rainfall brought by monsoon and tropical cyclones. Economically, Guangdong is the most prosperous province whose Gross Domestic Product (GDP) ranks the first among provinces in China. Guangdong, Hong Kong, and Macao composes a Greater Bay Area whose GDP exceeds 1.24 trillion USD, twice of San Francisco Bay Area. The landfall of severe typhoons, typhoon-related extreme precipitation events and storm surges usually lead to tremendously direct loss of property and lives. The economic loss in Guangdong province contributes an important share of annual cost in China due to typhoons [42]. Massive typhoons can result in employment increase and a periodic features to typhoons' impacts on employment [43]. The tropical cyclone strikes would cause significant (negative) impact on the economic in Guangdong province [44].

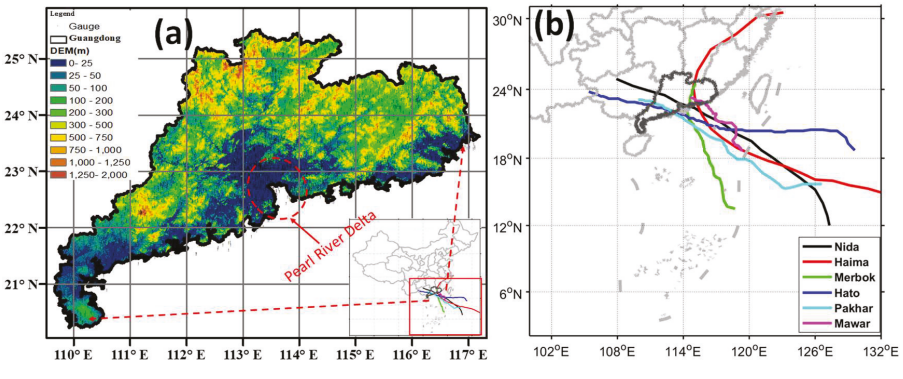


Figure 1. (**a**) Terrain and gauge distribution over Guangdong province in southern China. (**b**) Tracks of six typhoons. The typhoon paths data can be found on the link: http://tcdata.typhoon.org.cn/zjljsjj_zlhq.html.

2.2. Data

Observations from a dense gauge network composed of 2449 rain gauges are used as reference (Figure 1). To investigate how well the IMERG captured the spatial pattern of precipitation during six typhoon events, the hourly gauge observations were interpolated into regular gridded analysis (hereafter referred to as "GGKRIG") using a Kriging interpolation algorithm embedded in Version 8.2 Interactive Data Language (IDL). The spherical covariance function was used to model the variogram function in the Kriging interpolation algorithm. It needs to be kept in mind that the interpolation process may add uncertainties to the gridded gauge analysis, but these uncertainties were not considered in this study. Besides, the average rain gauge measurements within an individual pixel of IMERG data are used to compare against the measurements of IMERG pixel by pixel in order to mitigate the evaluation uncertainty raised by the inhomogeneity of precipitation that may cause a non-uniform beam filling (NUBF) problem for the space-born instruments onboard satellites [45].

The publically available latest version of IMERG algorithm is V5B. The V5B IMERG Final Run product was released on 30 November 2017, and the Early and Late Runs were released on 1 December 2017. The IMERG algorithm is a unified U.S. QPE algorithm that takes advantage of three U.S. satellite-based precipitation algorithms, i.e., PERSIANN-CCS [8], CMORPH [9], and the TMPA [10,11]. It integrates precipitation measurements derived from a numbers of PMW and IR sensors onboard the satellites to produce the quasi-global satellite-only QPE products provided in the Early and Late Runs, and combines the gauge analysis to yield gauge-corrected post research product in the Final Run [46]. The IMERG system runs twice in near-real-time to produce Early (Late) Run multi-satellite products after ~4 h (12 h) observation time. Once monthly gauge analysis is received in around 2.5 months after the observation month, the Final Run product will be produced by adjusting the half-hourly multi-satellite product to the monthly gauge analysis. Two kinds of gauge analysis products were used in the adjustment, i.e., the GPCC V7 full Data Analysis from 1998 to 2013 and the GPCC V5 Monitoring Product from 2014 to present. The adjustment approach is the legacy of the gauge calibration algorithm used in TMPA for infusing monthly gauge information into the fine-scale precipitation estimates [11]. Specifically, the ratio of monthly satellite-gauge field to multi-satellite-only field is used to multiply half hourly multi-satellite-only estimates to produce gauge-corrected half hourly QPE products. The differences between the Early Run and the Final Run algorithms include: (1) the Early Run only has forward propagation primarily for extrapolation, while the Final Run has both forward and backward propagation that allows interpolation; (2) the Final Run has ~2.5 months of latency that allows lagging data to be used in the Final run, even if these were not available for the Early Run; (3) the Early Run adopts trailing approach of rotation calibration of PMW with GPM Combined Radar-Radiometer (CORRA, using GMI and DPR), while the Final Run utilizes the centered approach; (4) surface pressure, surface temperature, and related humidity provided by Japan Meteorological Agency (JMA) forecasts are used in Early Run while those provided by European Centre for Medium-range Weather Forecasting (ECMWF) analysis are used in the Final Run; (5) the Early Run is calibrated to the Final Run with climatological coefficients that vary by month and location, while the Final Run is adjusted to the monthly gauge analysis following the TMPA approach; (6) different error functions for the individual satellite estimates are calculated using data over a three-month period, the trailing for the Early Run but centered on the target month for the Final Run. More details can be found in the IMERG Technical Documentation [47] and the release notes in V5B IMERG Final Run Release Notes at https://pmm.nasa.gov/sites/default/files/document_files/IMERG_FinalRun_V05_release_notes-rev3.pdf.

In this study, the Early Run satellite-only IMERG products (IMERG_ERUncal, hereafter) and the Final Run satellite-only and gauge-corrected IMERG products (IMERG_FRUncal and IMERG_FRCal, respectively, hereafter) are evaluated by comparing against the gauge observations. For fair comparison, the half hourly IMERG products were accumulated to hourly rainfall to match the temporary resolution of gauge observations. The readers need to keep in mind that the change of temporal resolution from half hourly to hourly may affect the accuracy of satellite data. Thus the results in this study are of

hourly scale. Additionally, the gauges used in this study may not be completely independent from the gauges used in IMERG, because there are approximately 25 out of the total 2449 gauges used in the GPCC.

2.3. Statistics Metrics

To quantitatively assess the performance of IMERG products, several commonly used skill scores are used for quantitative analysis. These skill scores include Bias, Relative Bias (RB), Root-Mean-Squared Error (RMSE), Correlation Coefficient (CC), Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), Hit Bias (HB), Miss Bias (MB), False Bias (FB), standard deviation (SD) of difference (SDD) between IMERG and gauge observations, relative SD (RSD), agreement index (AI) (Gebregiorgis et al., 2018). The Bias is the deviation of IMERG from gauge observations. The RB is dimensionless and used to measure the deviation of IMERG from gauge observations in percentage when it is multiplied by 100. The POD, CSI and FAR can be computed based on the number of hits (H), false alarms (F), and misses (M) shown in the contingency Table 2. The POD, CSI and FAR are used to measure IMERG' contingency scores. The AI measures the magnitude of agreement between IMERG and gauge observations. Additionally, HB, MB and FB are defined as the components of bias between gauge and IMERG. More detailed definition of these metrics can be found in Table 3.

Table 2. Contingency table of gauge and satellite estimates.

	Gauge ≥ Threshold	**Gauge < Threshold**
Satellite ≥ Threshold	Hits (H)	False alarms (F)
Satellite < Threshold	Misses (M)	Correct negative (C)

Table 3. Definition of statistic metrics to compare IMERG (S) QPE and gauge (G) observed QPE.

Metrics	**Equations**	**Perfect Value**
CC	$CC = \frac{Cov(S,G)}{\sigma_S \sigma_G}$	1
Bias	$Bias = \sum(S - G)$	0
RB (%)	$RB = \frac{\sum(S-G)}{\sum G} \times 100\%$	0
HB	$HB = \sum(S(S > 0 \,\&\, G > 0) - G(G > 0))$	0
MB	$MB = \sum S(S = 0 \,\&\, G > 0)$	0
FB	$FB = \sum S(S > 0 \,\&\, G = 0)$	0
POD	$POD = \frac{H}{H+M}$	1
CSI	$CSI = \frac{H}{H+M+F}$	1
FAR	$FAR = \frac{F}{H+F}$	0
SDD	$SD = \frac{\sum_{i=1}^{N}(x_i - \overline{x})}{N-1}$	0
RSD	$RSD = \frac{\sigma(S_i - G_i)}{\overline{G}}$	0
AI	$AI = 1 - \frac{\sum_{i=1}^{N}(S_i - G_i)}{\sum_{i=1}^{N}\left(\lvert S_i - \overline{G}\rvert + \lvert G_i - \overline{G}\rvert\right)^2}$	1

Note: N denotes the number of samples. In the formula of CC, "Cov()" represents the covariance, and "σ" the standard deviation. In the formula of SDD, $\overline{x}$ denotes the average of differences between IMERG and gauge observed QPE, x_i represents the difference between IMERG and gauge observed precipitation for a grid where there is at least one valid gauge observation exists at ith hour.

3. Results

3.1. Spatial Analysis

Figure 2 shows the spatial distributions of accumulated rainfall from GGKRIG, IMERG_ERUncal, IMERG_FRUncal and IMERG_FRCal during six typhoon events. Spatially, GGKRIG shows pronouncedly detailed texture alignment of precipitation while the precipitation of IMERG products varies smoothly from low precipitation area to intensive precipitation area. Overall, all three

IMERG products generally capture the spatial pattern of the precipitation depicted by GGKRIG over Guangdong province, but fail to capture some local orographic maxima indicated by red letter A in Figure 2a, B in Figure 2i, C in Figure 2m, D in Figure 2q. Compared with the typhoon tracks and the terrains as shown in Figure 1, this could be explained as follows. Firstly, warm rain accretion process dominates typhoon rainfall after landfall in China [48]. Secondly, the tremendous wet air masses brought by typhoons interacted with the local mountains in the upward side to produce shallow orographic rainfall. This type of rainfall system tends to be underestimated by PWM algorithm due to weak ice scattering signature [49]. As for IMERG_ERUncal, it shows similar spatial patterns and a little lower precipitation intensities when compared to IMERG_FRCal. After careful investigation, it is found that the Early Run and the Final Run IMERG products contain the same sources of PMW sensors, identical high quality (HQ, i.e., PMW-only) precipitation and IR-only precipitation, but have different IR Kalman filter weights during the six typhoon-related precipitation events. This indicates that the calibration of PMW with CORRA and the auxiliary data like JMA and ECMWF have no impact on the PMW-only and IR-only precipitation. Therefore, the differences between IMERG_ERUncal and IMERG_FRUncal are likely due to the fact that backward propagation algorithm was not used in IMERG_ERUncal and thus no weighting average between two consecutive microwave overpasses is applied. In contrast, IMERG_FRUncal utilizes both backward and forward propagation components at the expense of a little larger data latency (~8 h). This suggests that location and intensity of the precipitation events could be more accurately captured with additional measurements from the next microwave overpass. It is noted that IMERG_FRCal demonstrates lower accumulated rainfall than the other two IMERG satellite-only QPE products, especially during typhoon Pakhar and Mawar. Previous studies by [25,50,51] report that the gauge correction approach used in TMPA often fails to produce compatible rainfall intensities with local gauge observations during extreme precipitation events, indicating that the gauge correction approach used in IMERG still suffers from similar limitations and shortcomings. Similar findings about this gauge correction algorithm over complex topography during heavy rainfall events can be found in studies by Dinku et al. and other literatures [25,27,50,52].

Quantitatively, the scatter plots in Figure 3 show that both IMERG_ERUncal and IMERG_FRUncal tend to overestimate the accumulated precipitation except for typhoon Merbok and Mawar. The statistic indexes indicate that IMERG products, either Early Run or Final Run, have different behaviors. This is likely due to complicated typhoon structures, atmospheric structures, different paths, inherent spatio-temporal inhomogeneity of precipitation, and the complicated terrains in Guangdong province. Generally, IMERG_FRUncal has similar but better performance than IMERG_ERUncal with higher CC (0.59~0.82), smaller RMSE (44.83 mm~33.95 mm), and smaller absolute RB (−17.34%~17.33%). In particular, during typhoon Haima, IMERG_FRUncal shows better skills in estimating rainfall with distinctly higher CC (0.77 vs. 0.72), lower RB (15.32% vs. 30.97%), and smaller RMSE (37.34 mm vs. 42.33 mm). This confirms the positive merits of backward propagation algorithm used in IMERG Final Run process. For the gauge-corrected IMERG_FRCal, it demonstrates similar skills of CC and RMSE with IMERG_FRUncal, but generally tends to underestimate the storm-accumulated precipitation with RB ranging from −7.21% to 30.58%. It is noted that the absolute RBs of IMERG_FRCal are overall smaller than IMERG_ERUncal and IMERG_FRUncal, suggesting the effectiveness of gauge-correction in reducing deviation of satellite-only rainfall estimates from the observations. Besides, this also confirms the limitation of gauge-correction algorithm during extreme precipitation events using monthly gauge analysis in IMERG Final Run as the aforementioned.

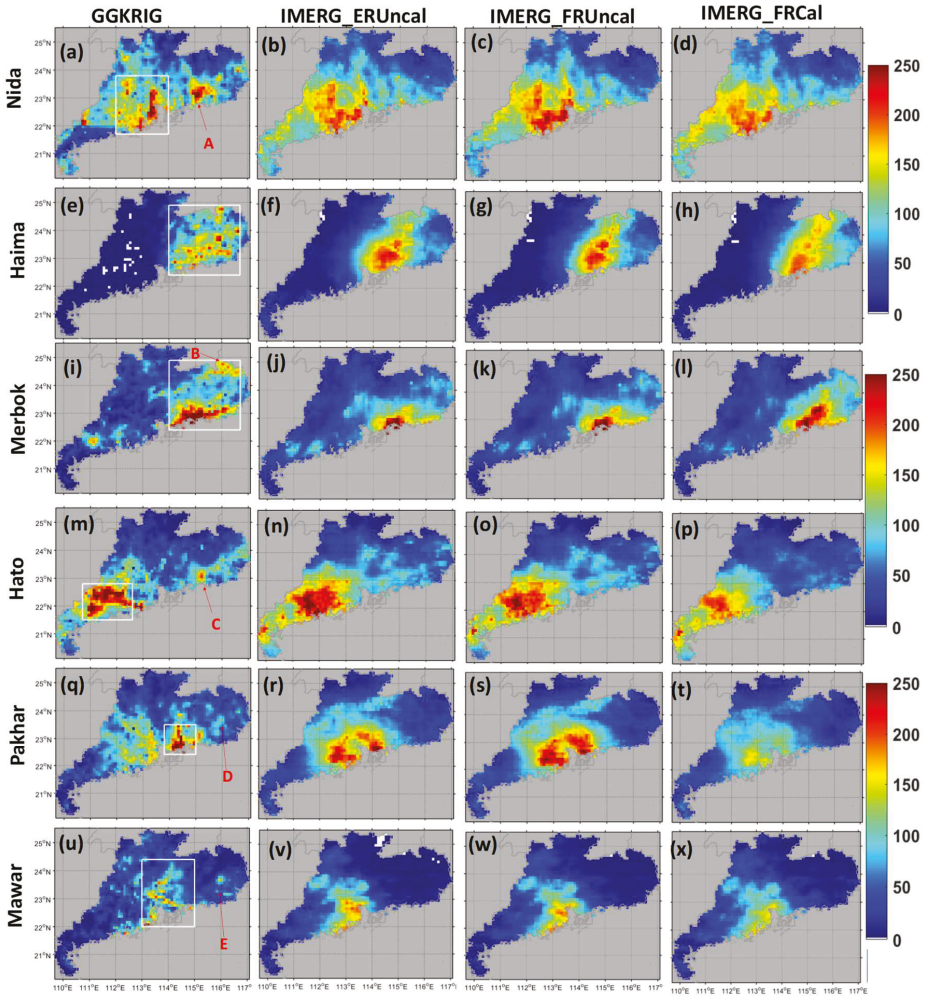


Figure 2. Spatial distribution of precipitation derived from gauge observations and IMERG Final Run products during typhoon (**a**–**d**) Nida, (**e**–**h**) Haima, and (**i**–**l**) Merbok, (**m**–**p**) Hato, (**q**–**t**) Pakhar, and (**u**–**x**) Mawar. The white rectangles in the first column denote the rainfall centers observed by gauge during different events.

To better understand the comprehensive performance of these three products, Figure 4 shows Taylor diagrams to visualize the statistics summary of how well the three IMERG products agree with the observations (GGKRIG) in terms of CC, RMSE, and standard deviation. The Taylor diagram was proposed by Taylor to provide a visual framework of summarizing multiple aspects of model performance related to reference (observations) [53]. The closer the model to the reference, the better the model. It is found that IMERG_FRCal has the best performance during typhoon Haima, Hato, Pakhar, and Mawar, and shows a little worse but close performance with IMERG_ERUnal and IMERG_FRUncal. IMERG_FRUncal performs better than IMERG_ERUncal in all typhoons except for typhoon Pakhar in which IMERG_FRUncal shows very close performance with IMERG_ERUncal.

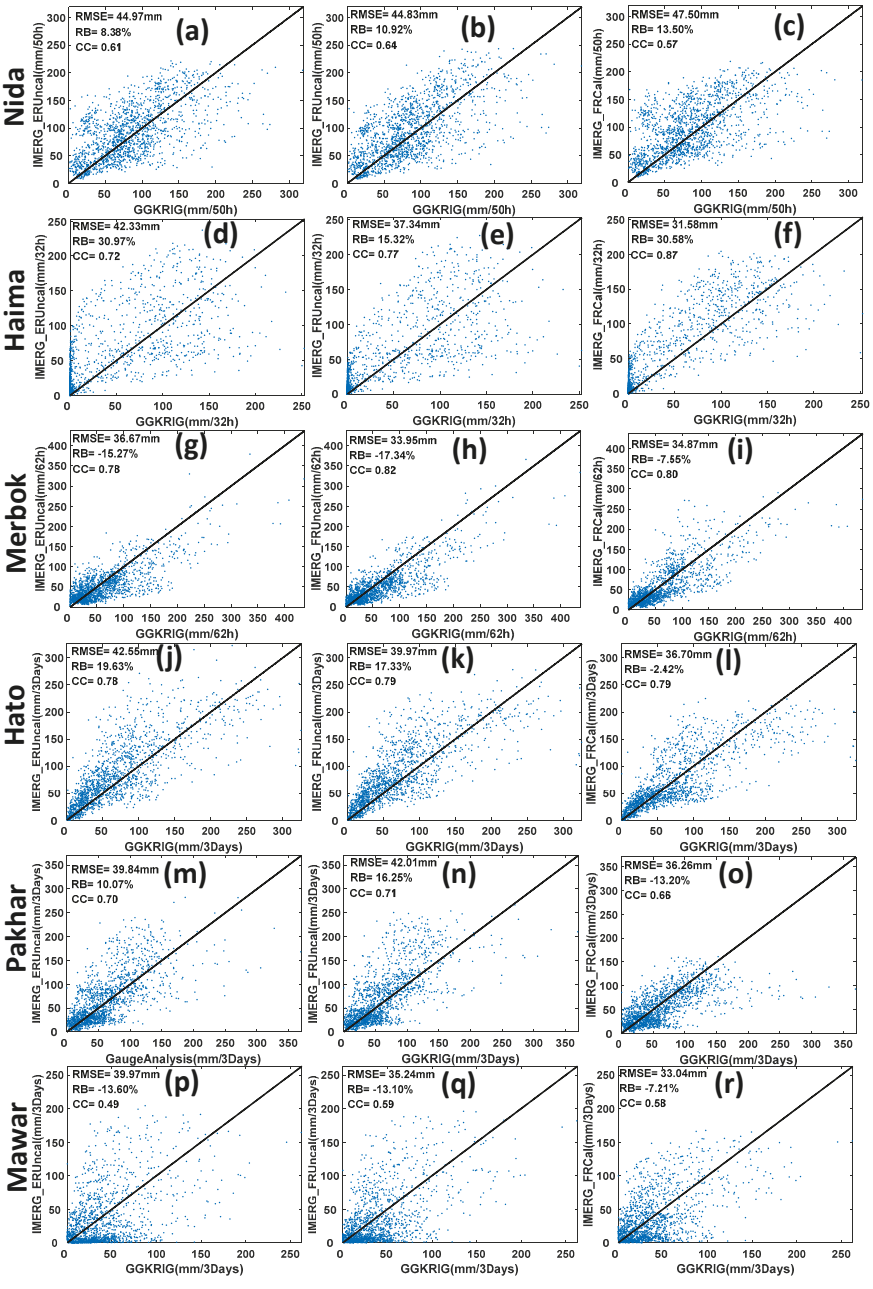


Figure 3. Scatter plots of gridded gauge observations versus IMERG Early Run (1st column), and Final Run (2nd and 3rd columns) products during typhoon (**a**–**c**) Nida, (**d**–**f**) Haima, (**g**–**i**) Merbok, (**j**–**l**) Hato, (**m**–**o**) Pakhar, and (**p**–**r**) Mawar. RMSE is short for root-mean-squared error, RB for relative bias, CC for correlation coefficient.

To further investigate the error sources of IMERG products, Figure 5 shows the spatial distribution of accumulated PMW-only estimates for IMERG (IMERG_PMW) and corresponding GGKRIG as well as the scatter plots between IMERG_PMW and GGKRIG during six typhoon-related rainfall events. It is noted that the IMERG_PMW significantly underestimated the precipitation, especially the rainfall center denoted by white rectangles (e.g., the center of rainfall during typhoon Hato shown in Figure 5j–k). Additionally, some local precipitation maxima related to orographic effects as arrowed by red letter A-I in Figure 5 are poorly detected and estimated by IMERG_PMW. In particular, IMERG_PMW shows no precipitation in southwestern Guangdong province while the gauges do during typhoon Haima. The scatter plots imply that the IMERG_PMW underestimated rainfall to a great extend from 42.97% to 69.71%, and poorly correlated with gauge observations with low CC less than 0.6 except typhoon Haima and Merbok. Figure 6 shows the hourly availability of PMW sensor observations over Guangdong province. It can be seen that most of the time (Nida: 54%; Haima: 43.75%; Merbok: 46.77%; Hato: 45.83%; Pakhar: 40.28%; Mawar: 45.83%) there is no precipitation derived from microwave sensor observations for IMERG for all six events. According to the IMERG algorithm, the IR-based precipitation is used to fill the area where there is no high quality PMW precipitation during these time slots. Thus it can be included that the lack of PMW sensor observations, the limitation of PMW-based precipitation as well as the shortcomings of IR-based precipitation are attributed to the performances of IMERG_ERUncal and IMERG_FRUncal during these six typhoon-related events.

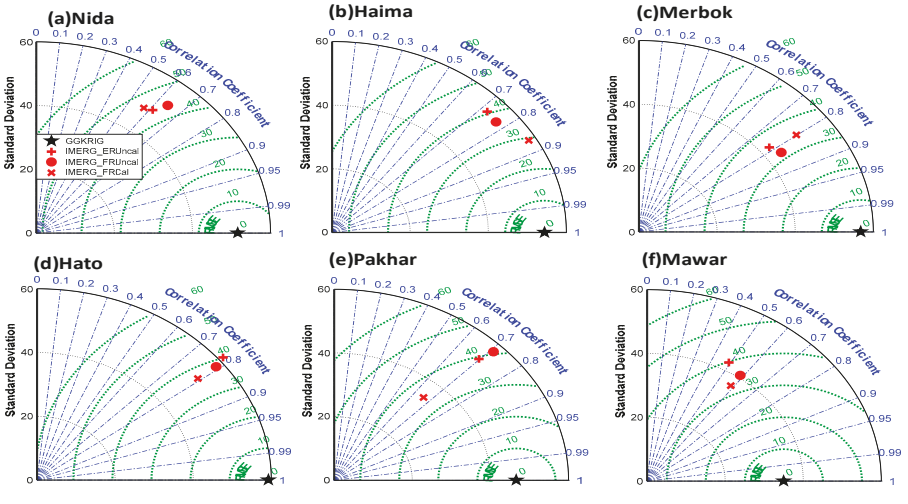


Figure 4. Taylor diagrams showing correlation coefficients, standard deviation (mm), and RMSE (mm) of storm-scale accumulated precipitation between the IMERG Early (Final) Run precipitation product and the reference data during typhoon (**a**) Nida, (**b**) Haima, (**c**) Merbok, (**d**) Hato, (**e**) Pakhar, (**f**) Mawar. The radical coordinate is the magnitude of satellite precipitation and standard deviation (mm), the concentric green semi-circles mean RMSE values, and the angular coordinate denotes the correlation coefficient.

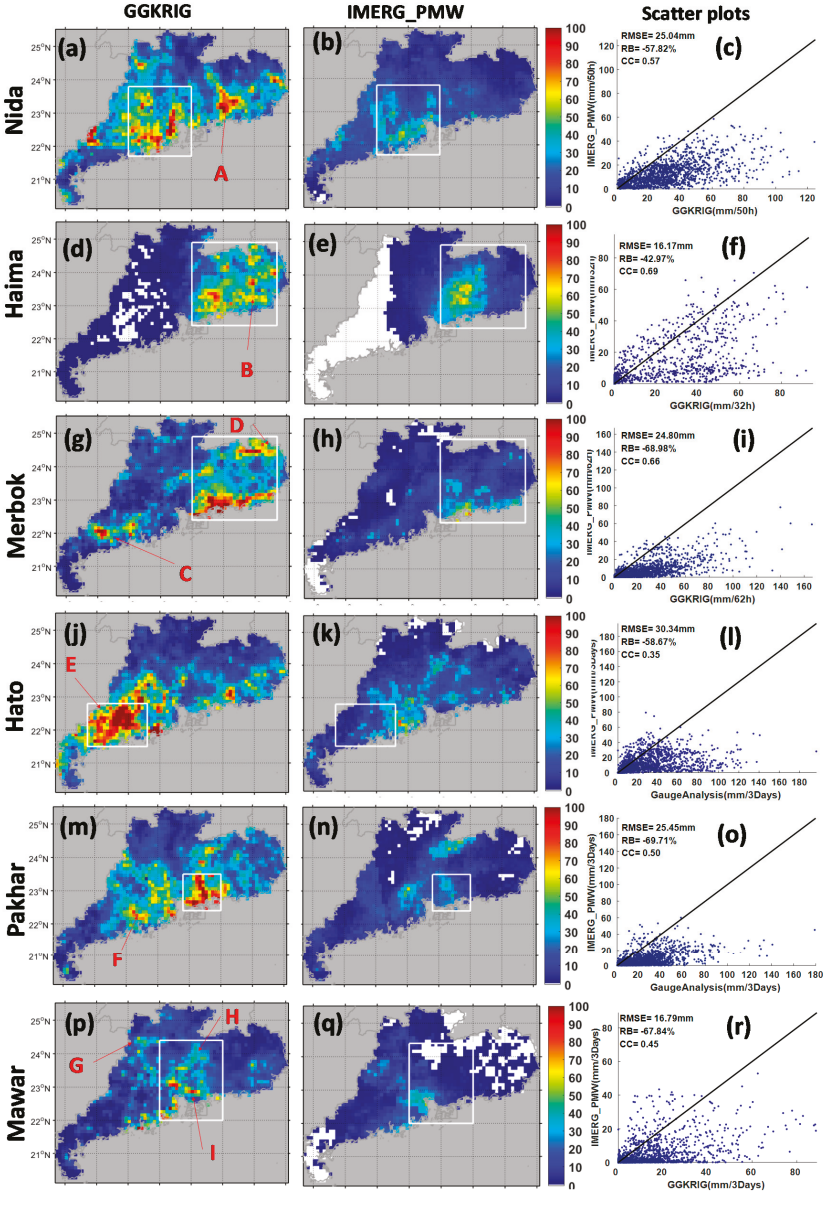


Figure 5. Spatial distribution of gridded gauge observations (1st column) versus satellite-based precipitation(2nd column), and the corresponding scatter plots (3rd column) under conditions in which passive microwave sensor observations were used by both IMERG Early and Final Run during typhoon (**a**–**c**) Nida, (**d**–**f**) Haima, (**g**–**i**) Merbok, (**j**–**l**) Hato, (**m**–**o**) Pakhar, and (**p**–**r**) Mawar. The red letters "**A**–**I**" denote the rain belts. Note that the subfigures in the 1st and 2nd columns are different from Figure 2 in which no filtering is applied to the precipitation dataset, whereas subfigures in the 1st and 2nd columns only illustrate the precipitation under conditions in which passive microwave sensor observations were both used by IMERG Early and Final Run.

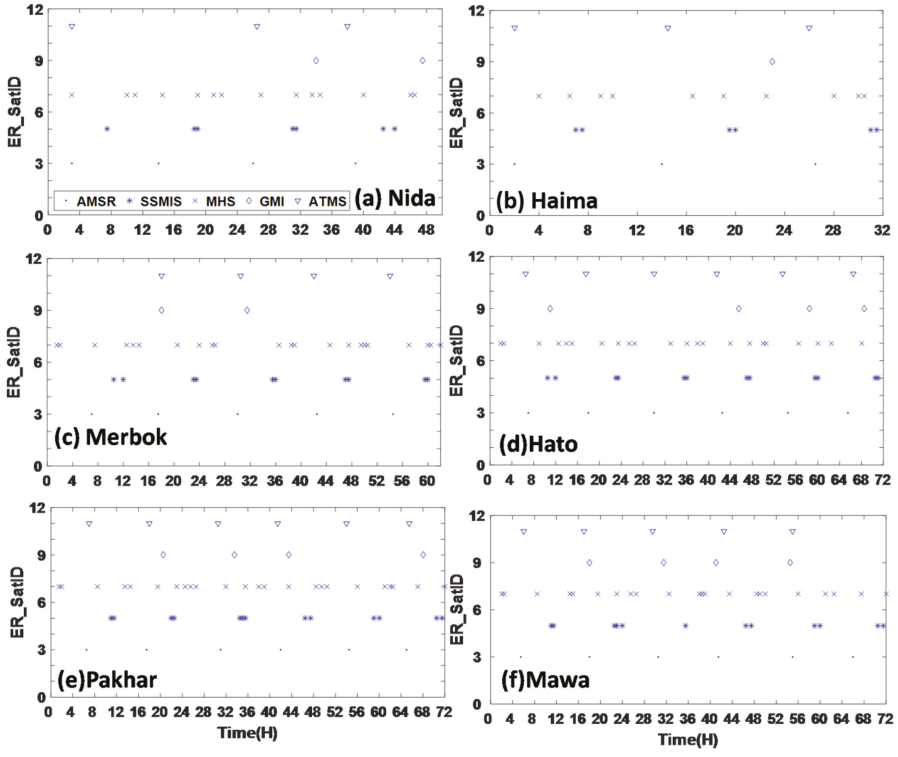


Figure 6. Hourly availability of passive microwave sensor observations over Guangdong province during typhoon (**a**) Nida, (**b**) Haima, (**c**) Merbok, (**d**) Hato, (**e**) Pakhar, and (**f**) Mawar. Here AMSR is short for Advanced Microwave Scanning Radiometer, SSMIS for Special Sensor Microwave Imager/Sounder, MHS for Microwave Humidity Sounder, GMI for GPM Microwave Imager, and ATMS for Advanced Technology Microwave Sounder.

Figure 7 illustrates the spatial distribution of IR-based precipitation used in the IMERG Early and Final runs. It is noted that the IR precipitation component used in the Final Run estimated much less than that in the Early Run especially during typhoon Mawar. As aforementioned, both the Early and Final Runs have identical PMW-only estimates and IR- based precipitation estimates, but have different IR Kalman filter weights. This indicates that: (1) different axillary data (JMA and ECMWF) and different approaches of calibration of PMW with CORRA did not exert any impact on the PMW-only and IR-only precipitation estimates; (2) the IR Kalman filter weights yielded by the forward propagation in Early Run are different from those produced by the forward and backward propagation in the Final run. In other words, IMERG_FRUncal integrated more PMW-only precipitation than IMERG_ERUncal during these typhoon-related precipitation events.

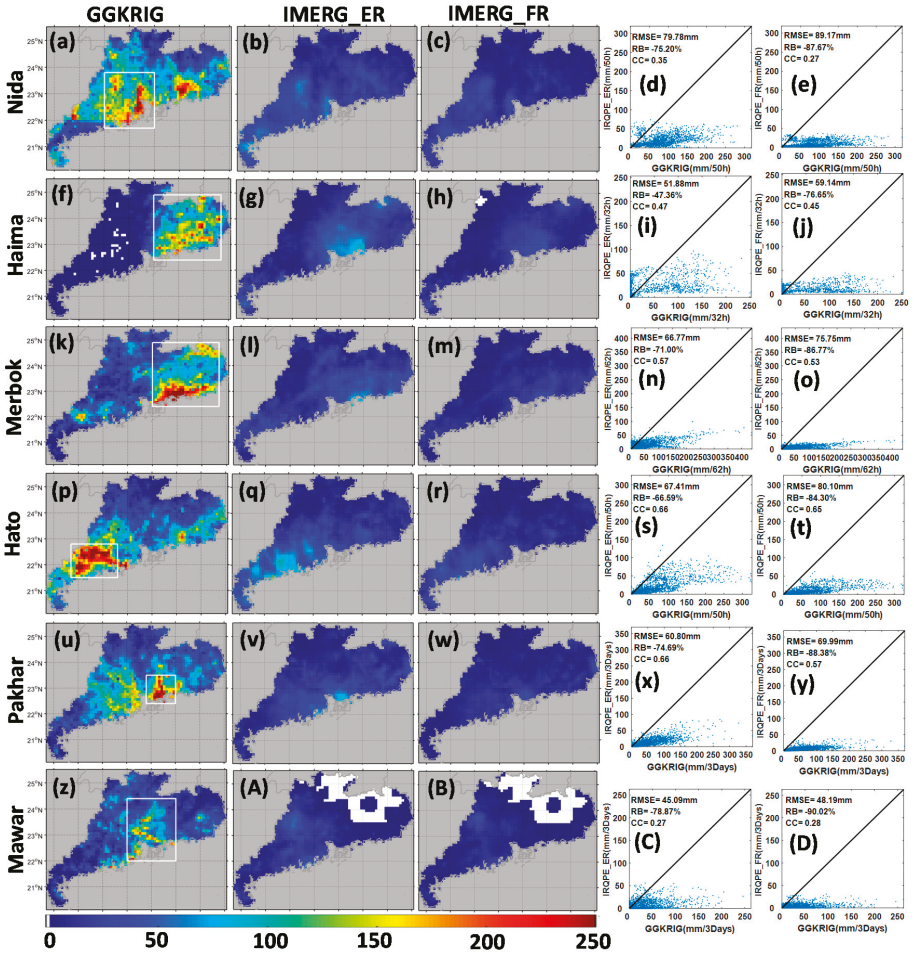


Figure 7. Spatial distribution of gridded gauge observations (1st column) versus satellite-based precipitation (2nd and 3rd columns), and the corresponding scatter plots (4th and 5th columns) under conditions in which infrared sensor observations were used by both IMERG Early (2nd column) and Final Run (3rd column) during typhoon (**a**–**e**) Nida, (**f**–**j**) Haima, (**k**–**o**) Merbok, (**p**–**t**) Hato, (**u**–**y**) Pakhar, and (**z**–**D**) Mawar. The white rectangles in the 1rd column denote the rainfall centers in six typhoon extreme events.

3.2. Hourly-Series Rainfall

Hydrologic cycle is greatly affected by temporal variation of precipitation. Flash flooding, landslides, and other hydro-related hazards are usually triggered by heavy precipitation over a short time period. Figures 8 and 9 show the plots of area-average hourly-series rainfall over Guangdong province and rainfall centers denoted in white rectangles in Figure 2 during six typhoon-related precipitation events. It needs to be kept in mind that only the grids that are overlapped with at least one gauge are taken into account to compute the hourly-series precipitation. It is noted that different typhoons have different number of precipitation peaks, which is likely attributed to the complicated typhoon structures and atmospheric environments. Generally, both the Early Run and

the Final Run IMERG product are able to capture the temporal variation of rainfall with gauge observations in all events with high CC > 0.7. Additionally, the temporal variation patterns of IMERG_ERUncal (IMERG_FRUncal, IMERG_FRCal) over the rainfall centers are similar to those over the whole Guangdong province. This indicates the precipitation over rainfall centers is the primary contribution to the total precipitation over Guangdong province. It is worth noticing that IMERG significantly overestimated the peak precipitation than gauge observations in all events except Typhoon Merbok. This is likely attributed to the fact that the half-hourly IMERG does not actually present half-hourly average rain rates, but rather a kind of quasi-instantaneous precipitation rate at some instantaneous moment during the half hour period when the satellite overpasses. Besides, it can be seen that IMERG_FRCal generally shows better agreement with the gauge observations during the rainfall peaks by significantly reducing the overestimations over rainfall centers especially during typhoon Hato and Pakhar. But IMERG_FRCal shows significant overestimation during typhoon Haima. This indicates that the gauge-correction algorithm with monthly gauge analysis in the Final Run still has limitations to conduct bias adjustment to the satellite-only QPE products.

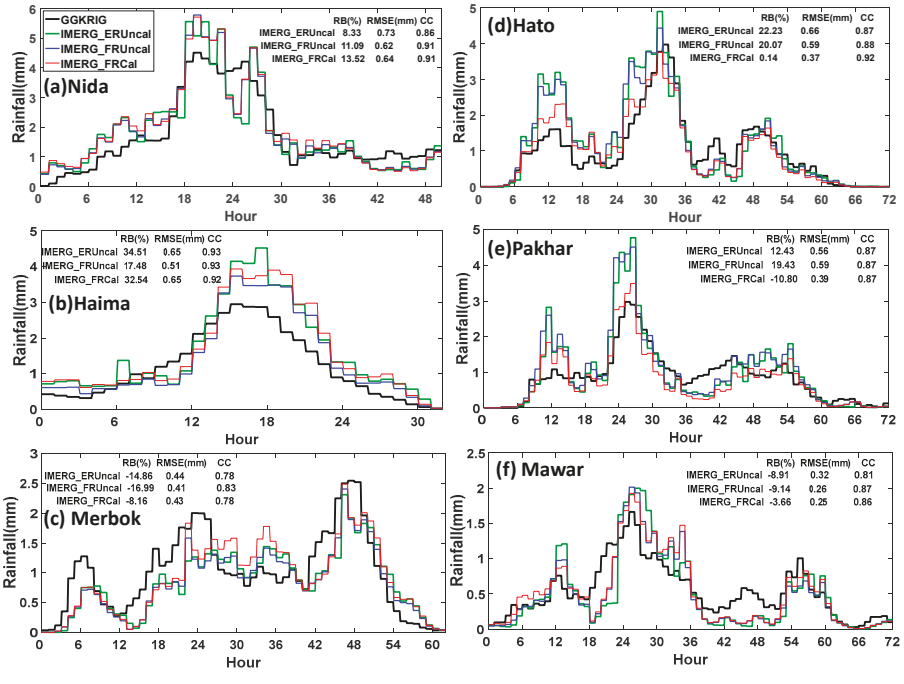


Figure 8. Area-average hourly rainfall over Guangdong province during typhoon (**a**) Nida, (**b**) Haima, (**c**) Merbok, (**d**) Hato, (**e**) Pakhar, and (**f**) Mawar.

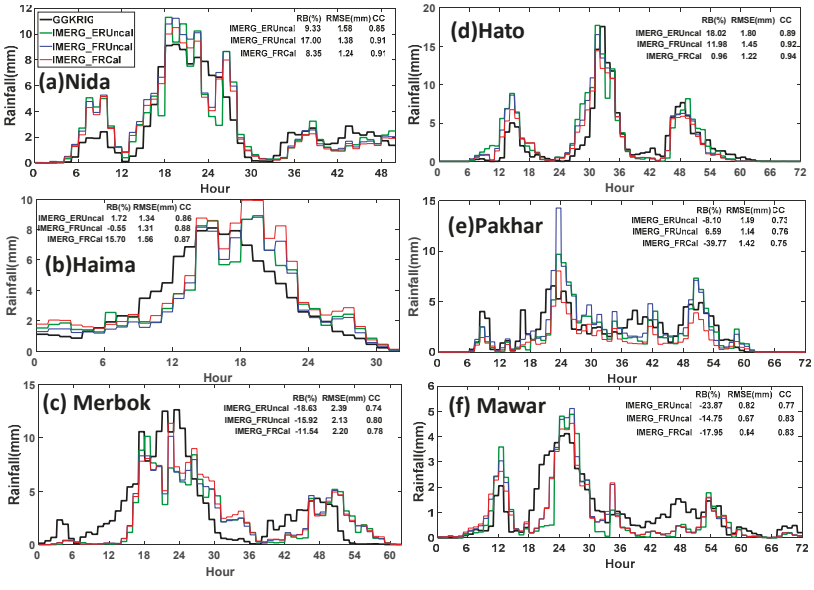


Figure 9. Area-mean hourly rainfall over rainfall centers denoted by white rectangles in Figure 2 during typhoon (**a**) Nida, (**b**) Haima, (**c**) Merbok, (**d**) Hato, (**e**) Pakhar, and (**f**) Mawar.

3.3. Probability Distribution by Occurrence

Figure 10 depicts the Probability Distribution as a Function (PDF) of rain rates by occurrence for gauge observations and the IMERG products during six heavy rainfall events. Only the grids that were overlapped with at least one gauge are used to compute PDF. Overall, all IMERG products underestimated the light precipitation (<0.6 mm/h) and overestimated the light to moderate precipitation in 0.6–6 mm/h. IMERG_ERUncal has close PDF with IMERG_FRUncal, and detected more light precipitation occurrences that IMERG_FRUncal. This is likely attributed to the fact that both forward and backward morphing algorithms were applied in the Final Run while only forward morphing algorithm was implemented in the Early Run.

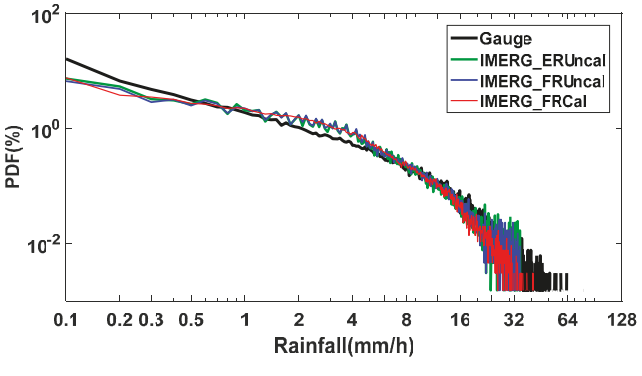


Figure 10. Probability distribution as a function of rain rates by occurrence with rainfall interval as 0.1 mm/h and log scale for x- and y-axis.

3.4. Analysis of Bias

The bias between IMERG products and the gauge observations is quantified in terms of arithmetic difference, and relative bias (difference) as shown in Figure 11. All of the IMERG products show underestimation and overestimation in different typhoon events. The Final Run gauge-corrected product IMERG_FRCal show smaller magnitude of RB ranging from −10.75% to 32.45%, and its counterpart satellite-only product IMERG_FRUncal has RB varying from −14.79% to 20.28%. The Early Run satellite-only product IMERG_ERUncal shows larger magnitude of RB ranging from −14.79% to 34.39%, indicating higher bias levels than the Final Run IMERG products.

To further investigate the sources of bias, the total bias was decomposed into three components: HB, MB, and FB. As shown in Figure 12, the FB is the primary component that contributes most to the total bias during all events. It is noted that the MB and FB are significantly reduced in the Final Run when compared to the Early Run for all events. This suggests the positive effectiveness of the backward morphing algorithm applied in the Final Run data processing. In addition, it can be seen that IMERG_FRCal tends to have negative HB, and both IMERG_FRCal and IMERG_FRUncal are inclined to have larger magnitude of HB than IMERG_FRCal. This implies that the success of forward and backward morphing algorithm in the Final Run may be at the expense of increasing HB.

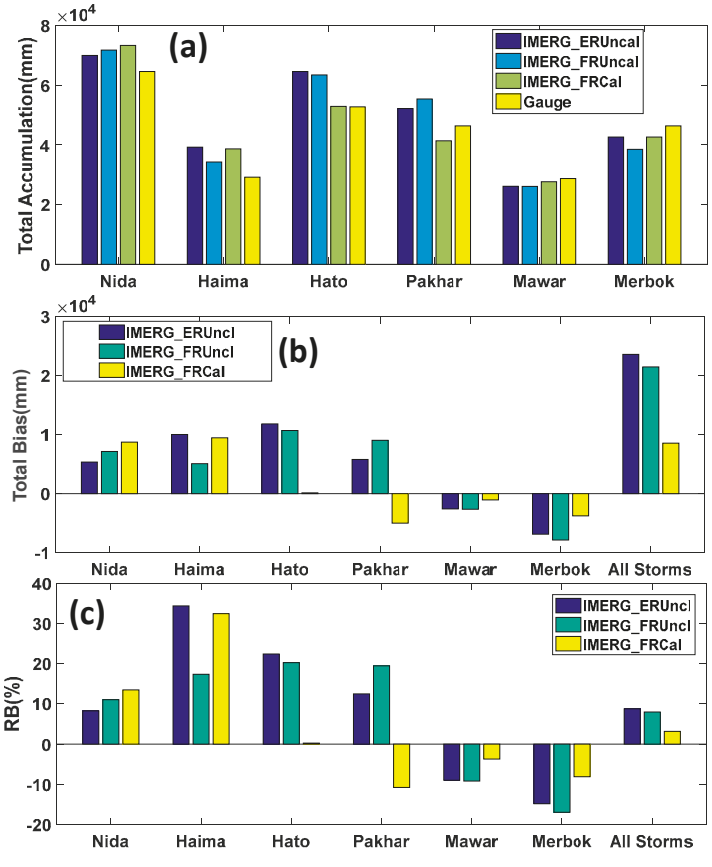


Figure 11. Summary statistics of the six typhoon extreme precipitation events: (**a**) total accumulated rainfall, (**b**) total accumulated bias, (**c**) relative bias.

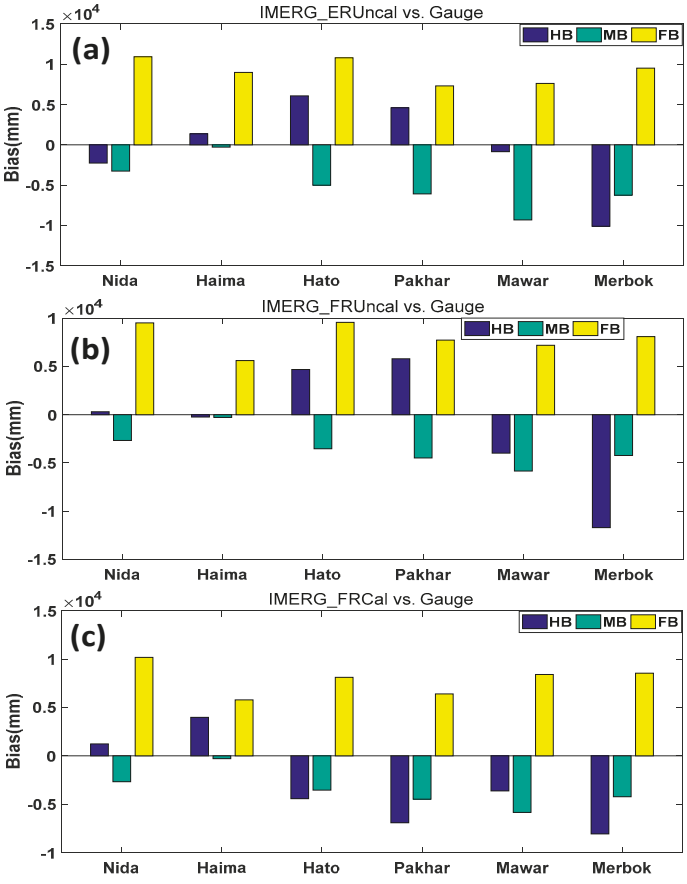


Figure 12. Decomposition of the total bias of the IMERG products for (**a**) IEMRG_ERUncal, (**b**) IEMRG_FRUncal, and (**c**) IEMRG_ERCal. Bias values are computed based on the pixel where there at least one gauge is located.

3.5. Contingency Analysis

Contingency analysis usually uses POD, FAR and CSI to quantify the performance of QPE products and precipitation forecast (QPF) products. Figure 13 shows the overall contingency scores for rainy event detection with 0 mm/h to separate rainy and non-rainy events. The IMERG_FRUncal and IMERG_FRCal have the same POD, FAR and CSI. This is because the IMERG_FRCal is obtained by multiplying IMERG_FRUncal with a scale factor which does not change IMERG_FRUncal's capability of detecting rainy events. Additionally, it is noted that both of them have higher POD and CSI but lower FAR than IMERG_ERUncal does. This once again confirms the positive effectiveness of forward and backward morphing algorithm in the Final Run on improving the rainy event detection.

For further contingency analysis, Figure 14 shows the contingency scores of different precipitation products as a function of different precipitation rates with interval of 1 mm/h. It is clearly shown that both IMERG_FRUncal and IMERG_FRCal have higher POD and CSI than IMERG_ERUncal when rain rates are less than 10 mm/h, and IMERG_ERUncal and IMERG_FRUncal demonstrate higher POD and CSI than IMERG_FRCal when rain rates are greater than 10 mm/h. Also, it can be found that IMERG_FRUncal and IMERG_FRCal demonstrate distinctly lower FAR than IMERG_ERUncal when

rain rates are less than 35 mm/h, and IMERG_FRCal has higher FAR than other two satellite-only products. This indicates these IMERG products have different skills in detecting precipitation. Also, this confirms that the gauge-correction with monthly gauge analysis in the Final Run processing leads to reducing high rain rates. Overall, the Final Run IMERG products outperform the Early Run IMERG product IMERG_ERUncal.

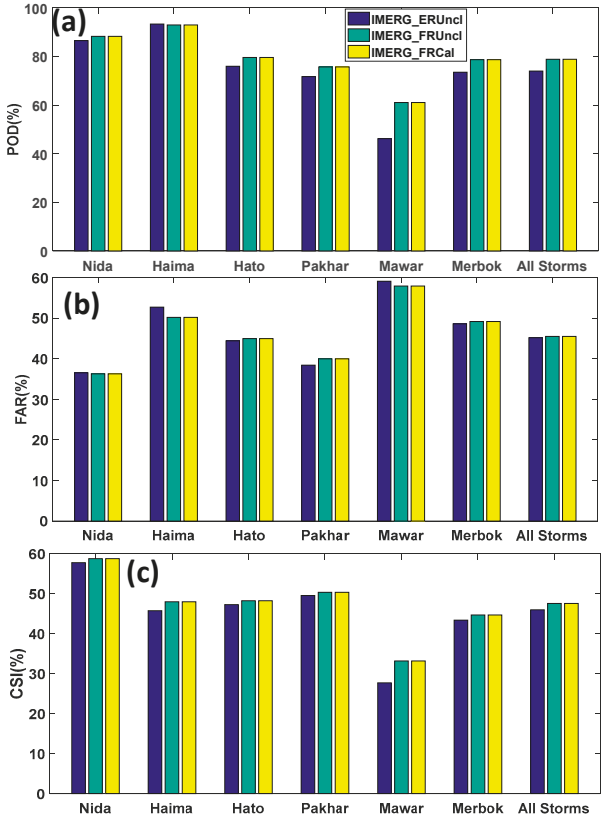


Figure 13. (**a**) Probabilities of detection (POD); (**b**) false alarm ratio (FAR); (**c**) critical success index (CSI). POD, FAR, and CSI were computed with 0 mm/h to separate rainy and non-rainy events.

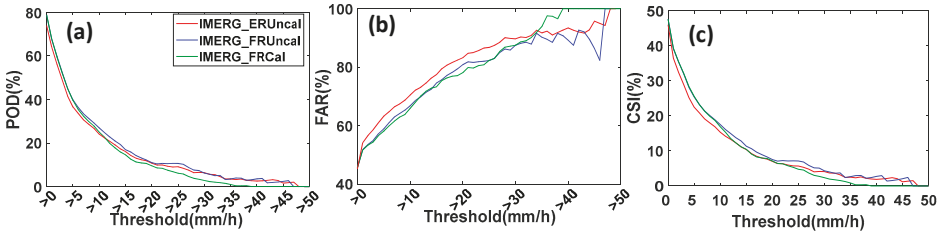


Figure 14. (**a**) Conditional probability of detection; (**b**) conditional probability of false detection; (**c**) conditional critical success index.

3.6. Analysis of Agreement and Disagreement Statistics

The agreement between IMERG and the gauge observations is quantified using Pearson's Correlation Coefficient (CC) and agreement Index (AI) (Figure 15). Both IMERG_FRUncal and IMERG_FRCal show close CC ranging from ~0.4 to 0.63, and moderate to high AI (0.6 to 0.75), and demonstrate pronouncedly higher CC and AI than IMERG_ERUncal. Low CC and AI values are reported during typhoon Mawar and Merbok. This re-confirms that the Final Run IMERG products benefit from the forward and backward morphing algorithm which can help improve precipitation estimation over the forward-only morphing algorithm applied in the Early Run.

The disagreement between IMERG and the gauge observations is assessed using standard deviation of differences (SDD) between IEMRG and gauge observations and relative SDD (RSD) (Figure 15c,d). It is noted that IMERG_ERUncal has evidently higher SDD than the Final Run IMERG products whose SDD are close to each other for most of events except Hato and Pakhar. Overall, IMERG_FRCal shows a little lower SDD than IMERG_FRUncal except Nida and Merbok. Regarding of RSD, Figure 15d depicts that the range of RSD is 2 to 5 times of the average rain rate during each storm. Both of IMERG_FRUncal and IMERG_FRCal demonstrates close RSD to each other, and lower RSD than IMERG_ERUncal does. Generally, in terms of RSD, IMERG_FRCal performs best, IMERG_ERUncal is ranked second, and IMERG_ERUncal performs worst. This may be likely linked to the forward and backward morphing algorithm and the gauge-correction algorithm implemented in the Final Run.

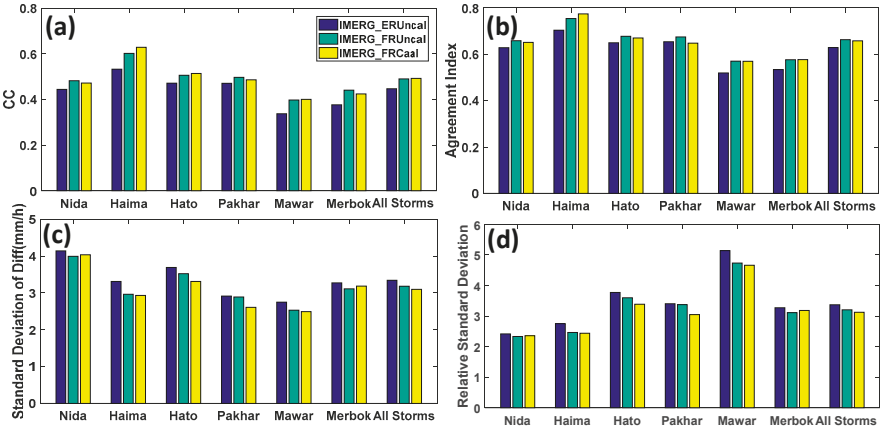


Figure 15. (**a**) Correlation coefficient; (**b**) standard deviation of differences (mm/h); (**c**) relative; (**d**) standard deviation of differences.

4. Summary and Conclusions

This study evaluated the V5 IMERG precipitation products during six typhoon extreme rainfall events aiming to provide the IMERG end users and algorithm developers with further insights on their accuracies when compared to the ground gauge observations. Both the Early Run and the Final Run IMERG products were selected for evaluations: the Early Run satellite-only IMERG_ERUncal, the Final Run satellite-only IMERG_FRUncal and gauge-corrected IMERG_FRCal. Due to the high spatio-temporal resolution which is expected to have great potentials of hydrologic utilization, the evaluation was conducted at the native spatial resolution of $0.1^\circ \times 0.1^\circ$ at hourly scale during six typhoon extreme precipitation events over Guangdong province in southern China. The observations from the dense gauge network composed of 2449 gauges were used as independent reference for evaluating the IMERG products. Spatial distribution of storm-accumulated precipitation, time series

rainfall, probability distribution, and contingency scores, bias decomposition, and the degree of agreement are analyzed in this study. The main conclusions based on the six typhoon-related storm cases are summarized as follows:

(1). IMERG products, either the Early Run or the Final Run, generally capture the spatiotemporal variability of precipitation with moderate to high CCs (0.6 to 0.87) for most of events except Mawar in which IMERG_ERUncal has low CC about 0.49 (Figure 2), and with RB ranging from −15.27% to 42.55%, and RMSE ranging from 31.58 mm to 47.50 mm (Figure 1).

(2). Overall, the IMERG products track reasonably well the temporal variation of the area-average hourly precipitation over Guangdong (with high CCs from 0.78 to 0.93) and rainfall centers (with high CCs from 0.74 to 0.94) during six typhoon-related events. All IMERG products tend to overestimate the peak rainfall especially IMERG_ERUncal overestimates the most. IMERG_FRUncal overestimated less than IMERG_ERUncal. IMERG_FRCal shows best agreement with gauge observations (Figures 7 and 8).

(3). All IMERG products underestimate the light precipitation (<0.6 mm/h) and overestimate the light to moderate precipitation in 0.6–6 mm/h (Figure 9).

(4). The FB is the primary component that contributes most to the total bias during all events. IMERG_FRCal tends to have negative HB, and both IMERG_FRCal and IMERG_FRUncal are inclined to have larger magnitude of HB than IMERG_FRCal does.

(5). IMERG_FRUncal and IMERG_FRCal have the same capability of detecting rainy events, showing higher POD and CSI, and lower FAR than IMERG_ERUncal. IMERG_FRUncal and IMERG_FRCal have higher POD and CSI than IMERG_ERUncal when rain rates are less than 10 mm/h. IMERG_ERUncal and IMERG_FRUncal demonstrate higher POD and CSI than IMERG_FRCal when rain rates are greater than 10 mm/h (Figure 13).

(6). The large magnitude of underestimation and overestimation of the heavy rainfall regions by satellite-only IMERG is primarily linked to the lack of microwave sensor observations. The limitations of IR-based and PMW-based rainfall estimation over complex terrains are attributed to the large bias of IMERG products when compared to gauge observations.

The V5B IMERG is state-of-the-art high resolution QPE products. It makes full use of advanced morphing algorithms and utilizes all available observations from IR and PMW sensors, the space-born dual-pol precipitation radar, and gauge observations over globe. This study suggests that IMERG products show promising performance in capturing the spatial and temporal patterns of very powerful typhoon extreme rainfall events like Nida, Haima, and Hato over Guangdong province in southern China, but still demonstrate large magnitude of overestimation for the satellite-only IMERG products from both the Early Run and the Final Run. The Final Run gauge-corrected IMERG_FRCal shows better performance than the satellite-only IMERG_ERUncal and IMERG_FRUncal, but it tends to underestimate the overall precipitation and still has considerable bias in estimating the hourly area-average precipitation. It is expected that the upcoming upgraded version of IMERG that uses daily gauge analysis for gauge correction will greatly improve the Final Run research product. Besides, these results are based only on six extreme typhoon-related precipitation storm, and more and further analysis are needed to investigate the sources of IMERG bias and its behaviors in different events. For example, there is a significant difference between the sample volumes of rain gauge and GPM sensors, and the one-dimension gauge measurement can't reflect the inhomogeneity of precipitation that causes a non-uniform beam filling (NUBF) problem for the space-born sensors [54]. Such NUBF problem could be further investigated with observations from local ground-based dual-pol weather radar and phase array radar network in Guangdong province. This work is undergoing.

Author Contributions: S.C. and J.H. conceived and designed the experiments; C.H. and A.Z. performed the experiments and analyzed the data; Z.L., L.X. and C.M. prepared the data and proofread the paper; X.T. and Z.Z. helped analysis the results; C.H. and S.C. wrote the paper.

Funding: This work was partially supported by "100 Top Talents Program" (74110-52601108) at Sun Yat-Sen University, Guangzhou, Guangdong, China, High-level Talents Training and Teacher Qualities and Skills Promotion Plan for Guangxi Colleges and Universities (8844), and by the National Natural Science Foundation of China (51579162, 41661021, 41875182), Guangxi Natural fund of Innovative Team Project (2016JJF15001).

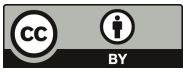
Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fischer, E.M.; Knutti, R. Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nat. Clim. Chang.* **2015**, *5*, 560. [CrossRef]
2. Rahmstorf, S.; Coumou, D. Increase of extreme events in a warming world. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17905–17909. [CrossRef] [PubMed]
3. Duffy, P.; Tebaldi, C. Increasing prevalence of extreme summer temperatures in the US. *Clim. Chang.* **2012**, *111*, 487–495. [CrossRef]
4. Kunkel, K.E.; Karl, T.R.; Brooks, H.; Kossin, J.; Lawrimore, J.H.; Arndt, D.; Bosart, L.; Changnon, D.; Cutter, S.L.; Doesken, N. Monitoring and understanding trends in extreme storms: State of knowledge. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 499–514. [CrossRef]
5. Rudeva, I.; Gulev, S.K. Climatology of Cyclone Size Characteristics and Their Changes during the Cyclone Life Cycle. *Mon. Weather Rev.* **2007**, *135*, 2568–2587. [CrossRef]
6. Clifton, H.E. Sedimentologic relevance of convulsive geologic events. *Geol. Soc. Am. Abstr. Progr.* **1985**, *17*, 1–6.
7. 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]
8. 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]
9. 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]
10. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Adler, R.F. Highlights of Version 7 TRMM Multi-satellite Precipitation Analysis (TMPA). In Proceedings of the 5th International Precipitation Working Group Workshop, Workshop Program and Proceedings, Hamburg, Germany, 11–15 October 2010; Klepp, C., Huffman, G., Eds.; Reports on Earth Sys. Sci. 100/2011; Max-Planck-Institut für Meteorologie: Hamburg, Germany, 2011; pp. 109–110.
11. 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]
12. Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Takayabu, Y.N.; Ushio, T.; Nakagawa, K.; Iwanami, 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]
13. Aonashi, K.; Awaka, J.; Hirose, M.; Kozu, T.; Kubota, T.; Liu, G.; Shige, S.; Kida, S.; Seto, S.; Takahashi, N. GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation. *J. Meteorol. Soc. Jpn.* **2009**, *87*, 119–136. [CrossRef]
14. Ushio, T.; Sasashige, K.; Kubota, T.; Shige, S.; Okamota, 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. Meterorol. Soc. Jpn. Ser. II* **2009**, *87*, 137–151. [CrossRef]
15. Shige, S.; Yamada, T.; Tsukiyama, T.; Kida, S. The GSMaP precipitation retrieval algorithm for microwave sounders. Part I: Over-ocean algorithm. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 5. [CrossRef]
16. Chen, S.; Hu, J.; Zhang, Z.; Behrangi, A.; Hong, Y.; Gebregiorgis, A.S.; Cao, J.; Hu, B.; Xue, X.; Zhang, X. Hydrologic Evaluation of the TRMM Multisatellite Precipitation Analysis Over Ganjiang Basin in Humid Southeastern China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4568–4580. [CrossRef]
17. Chen, S.; Hong, Y.; Cao, Q.; Kirstetter, P.-E.; Gourley, J.J.; Qi, Y.; Zhang, J.; Howard, K.; Hu, J.; Wang, J. Performance evaluation of radar and satellite rainfalls for Typhoon Morakot over Taiwan: Are remote-sensing products ready for gauge denial scenario of extreme events? *J. Hydrol.* **2013**, *506*, 4–13. [CrossRef]

18. 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**. [CrossRef]
19. Shen, Y.; Zhao, P.; Pan, Y.; Yu, J. A high spatiotemporal gauge-satellite merged precipitation analysis over China. *J. Geophys. Res. Atmos.* **2014**, *119*. [CrossRef]
20. Jiang, H.; Zipser, E.J. Contribution of tropical cyclones to the global precipitation from eight seasons of TRMM data: Regional, seasonal, and interannual variations. *J. Clim.* **2010**, *23*, 1526–1543. [CrossRef]
21. Gourley, J.J.; Hong, Y.; Flamig, Z.L.; Wang, J.; Vergara, H.; Anagnostou, E.N. Hydrologic Evaluation of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations on Ft. Cobb Basin, Oklahoma. *J. Hydrometeorol.* **2011**, *12*, 973–988. [CrossRef]
22. Ruin, I.; Lutoff, C.; Creton-Cazanave, L.; Anquetin, S.; Borga, M.; Chardonnel, S.; Creutin, J.; Gourley, J.; Gruntfest, E.; Nobert, S. Toward a Space-Time Framework for Integrated Water and Society Studies. *Bull. Am. Meteorol. Soc.* **2012**. [CrossRef]
23. 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]
24. Chen, S.; Liu, H.; You, Y.; Mullens, E.; Hu, J.; Yuan, Y.; Huang, M.; He, L.; Luo, Y.; Zeng, X. Evaluation of High-Resolution Precipitation Estimates from Satellites during July 2012 Beijing Flood Event Using Dense Rain Gauge Observations. *PLoS ONE* **2014**, *9*, e89681. [CrossRef] [PubMed]
25. Huang, Y.; Chen, S.; Cao, Q.; Hong, Y.; Wu, B.; Huang, M.; Qiao, L.; Zhang, Z.; Li, Z.; Li, W. Evaluation of Version-7 TRMM Multi-Satellite Precipitation Analysis Product during the Beijing Extreme Heavy Rainfall Event of 21 July 2012. *Water* **2013**, *6*, 32–44. [CrossRef]
26. Chen, S.; Hu, J.; Zhang, A.; Min, C.; Huang, C.; Liang, Z. Performance of near real-time Global Satellite Mapping of Precipitation estimates during heavy precipitation events over northern China. *Theor. Appl. Climatol.* **2018**, 1–15. [CrossRef]
27. 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]
28. 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**, *s176–s177*, 121–133. [CrossRef]
29. 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]
30. Ning, S.; Wang, J.; Jin, J.; Ishidaira, H. Assessment of the Latest GPM-Era High-Resolution Satellite Precipitation Products by Comparison with Observation Gauge Data over the Chinese Mainland. *Water* **2016**, *8*, 481. [CrossRef]
31. Ning, S.; Song, F.; Udmale, P.; Jin, J.; Thapa, B.R.; Ishidaira, H. Error Analysis and Evaluation of the Latest GSMap and IMERG Precipitation Products over Eastern China. *Adv. Meteorol.* **2017**, *2017*, 1–16. [CrossRef]
32. 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]
33. 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]
34. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]
35. 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]
36. Oliveira, R.; Maggioni, V.; Vila, D.; Porcacchia, L. Using Satellite Error Modeling to Improve GPM-Level 3 Rainfall Estimates over the Central Amazon Region. *Remote Sens.* **2018**, *10*, 336. [CrossRef]
37. 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]
38. Omranian, E.; Sharif, H.; Tavakoly, A. How well can global precipitation measurement (GPM) capture hurricanes? case study: Hurricane Harvey. *Remote Sens.* **2018**, *10*, 1150. [CrossRef]
39. 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]

40. Zhang, A.; Xiao, L.; Min, C.; Chen, S.; Kulie, M.; Huang, C.; Liang, Z. Evaluation of latest GPM-Era high-resolution satellite precipitation products during the May 2017 Guangdong extreme rainfall event. *Atmos. Res.* **2019**, *216*, 76–85. [CrossRef]
41. Khan, S.; Maggioni, V.; Kirstetter, P.E. Investigating the Potential of Using Satellite-Based Precipitation Radars as Reference for Evaluating Multisatellite Merged Products. *J. Geophys. Res. Atmos.* **2018**, *123*, 8646–8660. [CrossRef]
42. Elliott, R.J.R.; Strobl, E.; Sun, P. The local impact of typhoons on economic activity in China: A view from outer space. *J. Urban Econ.* **2015**, *88*, 50–66. [CrossRef]
43. Wu, X.; Zhou, L.; Guo, J.; Liu, H. Impacts of Typhoons on Local Labor Markets based on GMM: An Empirical Study of Guangdong Province, China. *Weather Clim. Soc.* **2017**, *9*, 255–266. [CrossRef]
44. Del Valle, A.; Elliott, R.J.R.; Strobl, E.; Tong, M. The Short-Term Economic Impact of Tropical Cyclones: Satellite Evidence from Guangdong Province. *Econ. Disasters Clim. Chang.* **2018**, *2*, 225–235. [CrossRef]
45. Durden, S.L.; Haddad, Z.; Kitiyakara, A.; Li, F. Effects of nonuniform beam filling on rainfall retrieval for the TRMM precipitation radar. *J. Atmos. Ocean. Technol.* **1998**, *15*, 635–646. [CrossRef]
46. 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). Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf (accessed on 10 January 2018).
47. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation*; Project, G., Ed.; NASA: Greenbelt, MD, USA, 2018.
48. 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]
49. Wen, L.; Zhao, K.; Chen, G.; Wang, M.; Zhou, B.; Huang, H.; Hu, D.; Lee, W.C.; Hu, H. Drop Size Distribution Characteristics of Seven Typhoons in China. *J. Geophys. Res. Atmos.* **2018**, 1–20. [CrossRef]
50. Kubota, T.; Ushio, T.; Shige, S.; Kida, S.; Kachi, M.; Okamoto, K.i. Verification of high-resolution satellite-based rainfall estimates around Japan using a gauge-calibrated ground-radar dataset. *J. Meteorol. Soc. Jpn.* **2009**, *87*, 203–222. [CrossRef]
51. Liao, L.; Meneghini, R. A Study on the Feasibility of Dual-Wavelength Radar for Identification of Hydrometeor Phases. *J. Appl. Meteorol. Climatol.* **2010**, *50*, 449–456. [CrossRef]
52. Dinku, T.; Connor, S.J.; Ceccato, P. *Comparison of CMORPH and TRMM-3B42 over Mountainous Regions of Africa and South America*; Springer: New York, NY, USA, 2010.
53. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]
54. Chen, H.; Chandrasekar, V. Validation of NASA's Global Precipitation Measurement mission with a high-resolution ground radar network. In Proceedings of the URSI Asia-Pacific Radio Science Conference (URSI AP-RASC), Seoul, Korea, 21–25 August 2016; pp. 836–839.



remote sensing

MDPI

Letter

Mesoscale Resolution Radar Data Assimilation Experiments with the Harmonie Model

Serguei Ivanov [1,*], Silas Michaelides [2] and Igor Ruban [1]

[1] Odessa State Environmental University (OSENU), 15, Lvovska Str., 65016 Odessa, Ukraine; iggru@i.ua
[2] The Cyprus Institute, 20, Konstantinou Kavafi Str., Aglantzia, 2121 Nicosia, Cyprus; s.michaelides@cyi.ac.cy
* Correspondence: svvivo@te.net.ua; Tel.: +38-050-631-3655

Received: 7 August 2018; Accepted: 7 September 2018; Published: 11 September 2018

Abstract: This study presents a pre-processing approach adopted for the radar reflectivity data assimilation and results of simulations with the Harmonie numerical weather prediction model. The proposed method creates a 3D regular grid in which a horizontal size of meshes coincides with the horizontal model resolution. This minimizes the representative error associated with the discrepancy between resolutions of informational sources. After such preprocessing, horizontal structure functions and their gradients for radar reflectivity maintain the sizes and shapes of precipitation patterns similar to those of the original data. The method shows an improvement of precipitation prediction within the radar location area in both the rain rates and spatial pattern presentation. It redistributes precipitable water with smoothed values over the common domain since the control runs show, among several sub-domains with increased and decreased values, correspondingly. It also reproduces the mesoscale belts and cell patterns of sizes from a few to ten kilometers in precipitation fields. With the assimilation of radar data, the model simulates larger water content in the middle troposphere within the layer from 1 km to 6 km with major variations at 2.5 km to 3 km. It also reproduces the mesoscale belt and cell patterns of precipitation fields.

Keywords: harmonie model; radar data assimilation; pre-processing; mesoscale precipitation patterns

1. Introduction

Precipitation plays an important role in both the water cycle and energy balance of the atmosphere. However, due to the high spatial and temporal variability of precipitation in the mesoscale, obtaining accurate quantitative precipitation estimates is still a "first-line frontier" task. Moreover, the Global Precipitation Measurement (GPM) mission has shown notable differences in estimations of precipitation obtained from various platforms especially for low rain rates [1]. Current rain-gauge networks mainly suffer from sparse distribution and limited coverage [2]. Although satellite-based infrared and visible data have high spatial resolution, relations between the radiance from the cloud and precipitation are indirect and non-unique [3]. Additionally, the evaluations of satellite quantitative precipitation estimates have some limitations in terms of the deficiency in the observations, the evaluation methodology, the selection of time windows for evaluation, and the short periods for evaluation [4]. Passive microwave instruments provide acceptable estimates of precipitation. Nevertheless, they exhibit low spatial and temporal resolution [5]. The advantage of the ground-based radar network in Europe is the high spatiotemporal resolution. This is the reason to make a choice in favor of such data for predicting mesoscale precipitation patterns and associated atmospheric variables.

The Operational Program for the Exchange of Weather Radar Information (OPERA) was launched for improving the harmonization of radars and their measurements [6]. The program provides three composite products such as instantaneous surface rain rate, instantaneous maximum reflectivity,

and one-hour rainfall accumulation. Gathered data includes three-dimensional volumes of reflectivity and radial wind. When assimilated in convection-permitting models, radar data could yield significant improvements in the representation of the spatial distribution of cloud and precipitation as well as rain rates [7–9]. This is achieved due to the fact that ground-based radars have a high spatial and temporal resolution and provide a three-dimensional volume of information. It is worth noting that radar data are heterogeneously distributed by providing abundant information near radar locations while such information becomes sparser with an increasing distance from the sensor. Further improvement of assimilation systems is seen in the optimization of pre-processing radar data in terms of homogenizing data coverage along with increasing confidence in them by averaging stochastic errors.

The objective of this study is to present a pre-processing approach adopted for the radar reflectivity data assimilation and the results of simulations with the Harmonie numerical weather prediction model. Through a case study, it will be shown that the method can improve the prediction of precipitation within the radar location area both in terms of the rain rates and spatial pattern presentation.

Even though radars measure the reflectivity and radial wind, this study focuses at the former parameter only. Quality control and other extra technical procedures are also omitted by assuming that they have been performed at preliminary steps during the initial processing of radar data. Further pre-processing is performed in different ways and may include various methods. Among these is the spatial screening of the raw data over some pixels in pursuit of both to reject abundant data and rid horizontal correlations in observation errors. The screening includes either filtering or smoothing procedures. The former is an easy way to routinely perform thinning of high-resolution radar data such as selection data of every predefined bin along a ray. Simple thinning of abundant radar observations throws a significant part of data out from the process, but it also affects the spectral distribution of sub-scales finer than a thinning parameter. An alternative to the thinning is the averaging of observations within a given box to create a new value located at an average position. This also allows us to average out random observation errors [10]. This method, which is called superobbing, is used for remote sensing observations such as atmospheric motion vectors and Doppler radar radial wind measurements [11]. With this approach, the observation (O) minus the background (B) differences (denoted as O–B) or innovations are averaged. For each observation, its model simulated analogue is computed and an O–B innovation is calculated. These innovations are then averaged and added to the model observation closest to the center of the superobbing cell to provide the super-observation. Optimization of the superobbing processing from dense raw data is the compromise between the above two factors including saving finer than average size scales and obtaining more stable estimates for the remaining scales by reducing stochastic errors [12]. Additionally, the problem of representativeness is necessary to be accounted for while comparing and assimilating data from sources with different spatial resolutions. Experience with assimilating high-resolution data shows that such data coverage may provide unsatisfactory results. In particular, tests with the Generalized Cross-Validation (GCV) method [13] using a simulated high-resolution data set at full resolution has led to a poorer analysis than a lower resolution data set that preceded it. The larger the nonconformity between the resolution of the observations network is, on the one hand, and that of the model, on the other hand, the larger the numerical impact on the result is or, equivalently, the larger the representativeness error is.

A modified method of pre-processing is presented in Section 2 in which the model and data used in this study are elaborated on. Results of simulations with the proposed pre-processing approach are presented in Section 3. Section 4 presents the discussion and concluding remarks.

2. Materials and Methods

2.1. Pre-Processing of Radar Data

In this study, the data assimilation system in the convection-permitting Harmonie model is further developed by involving radar reflectivity measurements. The focus is on optimizing the inner parameters of the pre-processing procedure rather than a data assimilation algorithm itself.

In pursuing the compatibility between the model resolution and smoothed radar observation density, the "cube-smoothing" approach is proposed. It seems obvious to optimally feed the model with observations at the same resolution as the model grid. Similar approaches at the assimilating step are explored for simulated reflectivity in a 1D + 3D variational assimilation system by feeding vertical profiles of reflectivity on a regular grid [14], an Ensemble Kalman Filter Analysis (EnKF) [15], and a Hybrid EnKF-3Dvar algorithm [16]. The earlier method performs backward retrieving of radar reflectivity from atmospheric microphysics variables for further assimilating measurements. This approach is operationally used in the model and pursuits the aim to globally assimilate radar data by doing it with a coarser resolution than explored in the regional model. The two latter approaches were developed within the Advanced Regional Prediction System (ARPS) and verified with the use of an Observing System Simulation Experiment (OSSE). However, as previously mentioned [13], the involvement of simulated high-resolution data sets in assimilating can, under certain circumstances, lead to an improper analysis.

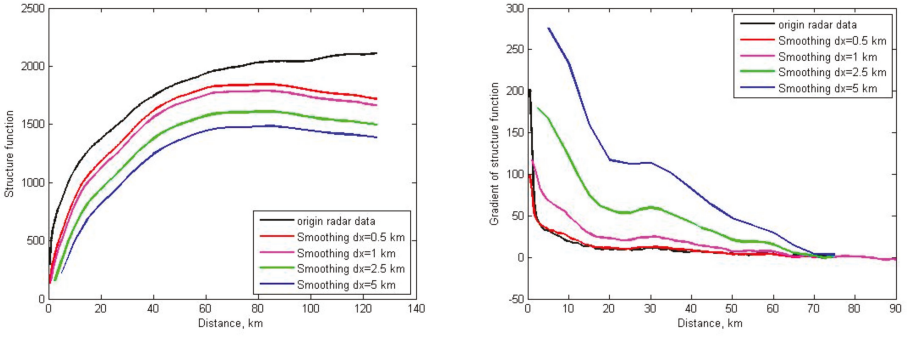
The proposed pre-processing method for smoothing radar reflectivity data imposes a minor additional computational cost, which can be afforded before the model is run and tuned for a specific domain. The method first creates a 3D regular grid in which a horizontal size of meshes coincides with a horizontal model resolution. This should minimize the error associated with the discrepancy between resolutions of informational sources [17] while the impact from filtering out sub-grid scales may occur. This is seen from comparing structure functions for original radar reflectivity data and those calculated after pre-processing (Figure 1). A Structure function as an alternative statistic to view variations in time series or fields is often used in meteorology [18,19] and is denoted by the equation below.

$$D_{AA}(L) = \frac{1}{N}\sum_{k=0}^{N-1}\left[A_k - A_{k+j}\right]^2 \quad (1)$$

where $L = r = j\Delta r$ represents the spatial separation, r, between the two measurements. The advantage of the structure function compared to other statistics is in its low dependence on variations between ensemble members as well as the possibility for obtaining stable estimates on the basis of a single limited length series containing systematic errors. For zero separation distance, the structure function is identical to zero and increases with the increasing separation until it reaches its maximum value, i.e., when it saturates. Such a distance implies lack of connectivity in a field and serves as a criterion for dominating scales in atmospheric processes. Regarding the slow changing structure functions for most atmospheric variables, the gradient of the function allows for the identification of a point on the curve, which corresponds to this criterion.

To estimate horizontal size and shape of precipitation patterns in the original data and after pre-processing with various smoothing values, the structure functions for radar reflectivity have been obtained. Only measurements from lower bins have been taken into account on the assumption that it provides horizontal scales in precipitation fields. Figure 1 shows that the saturation of the structure function starts at a few kilometer distance. Smoothing with 0.5 km yields very similar results both in the value and behaviors of the original data as the gradient. The latter allows us to clearly specify changes in the function behavior. In particular, saturation of the structure functions occurs at the shortest lag of about 1 km, which confirms the dominant role of mesoscales. Smoothing with the 2.5 and 5 km shows weaker gradients at the smallest resolved scales but also background (secondary) phenomena appear at the scales of about 20 to 30 km. This numerical (artificial) effect, which results from the inconsistency between filtering parameters and existent atmospheric patterns, is one of the reasons leading to a systematic phase error in the modeling of precipitation [20]. All structure functions completely saturate above the 60 km lag where the large-scale circulation dominates.

The drawback even for complete radar data sets is that they fill-in only a part of the physical or model spaces. For instance, the six Finish radars used in this study cover from 5% to about 20% of a model grid depending on the horizontal smoothing parameter and the vertical factor at the pre-processing step (Figure 2).



(**a**) (**b**)

Figure 1. (**a**) Structure functions and (**b**) their gradients of radar reflectivity measurements for original data and after pre-processing with various smoothing parameters.

This is a considerable difference between the OSSE approach and the real data sets. A vertical size for meshes can be chosen on the basis of several criteria. Among the most appropriate ones are those that correspond to a list of model layers with varying steps along the vertical coordinate to homogeneous filling-in meshes by radar measurements and to a regular vertical step. Regarding the irregularly distributed Finnish radar network, the preliminary estimates of filling-in 3D-meshes have been obtained for the whole domain in one batch (Figure 3). The use of the vertical step in accordance to the Harmonie model level list leads to minor filling-in of lower layers, which significantly increases. Otherwise, the regular but rather large vertical step of 500 m yields an opposite distribution of the data. So far, the equal filling-in approach seems to ensure the homogeneous saturation of the model grid, which was selected for further pre-processing. At least three direct impacts from the pre-processing smoothing are clear. Over-smoothing with a parameter larger than the model's resolution leads to underestimating the total amount of water content and the small decreases but intensive cells of precipitation as well as the incorrectly presenting spatial patterns of rain water in the atmosphere. To this end, these factors affect the initial conditions and disrupt the subsequent evolution of precipitation patterns. Smoothing with 0.5 km yields reflectivity distributions similar to the original (Figure 2a).

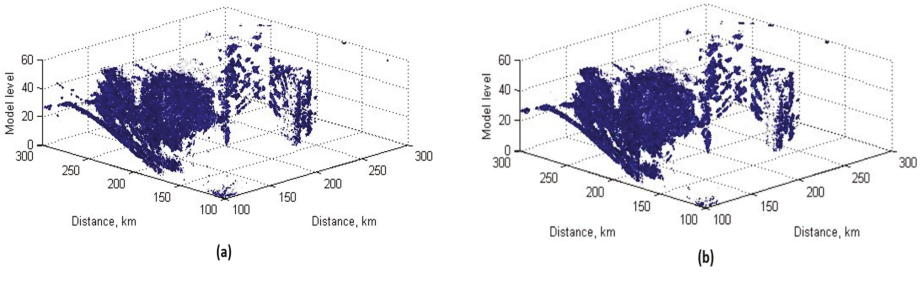


Figure 2. *Cont.*

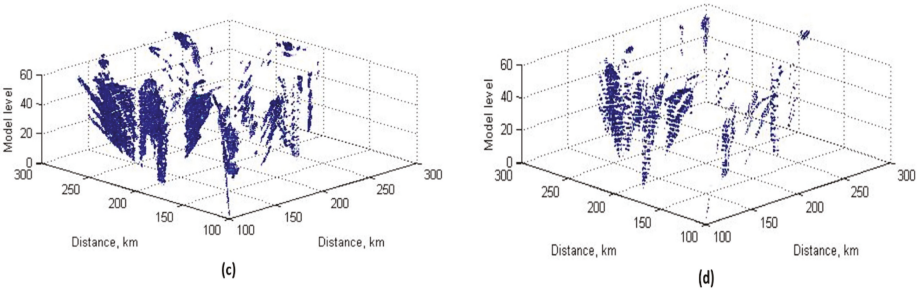


Figure 2. Filling-in the model grid by radar measurements depending on pre-processing smoothing parameters: (**a**) 0.5 km, (**b**) 1 km, (**c**), 2.5 km, and (**d**) 5 km.

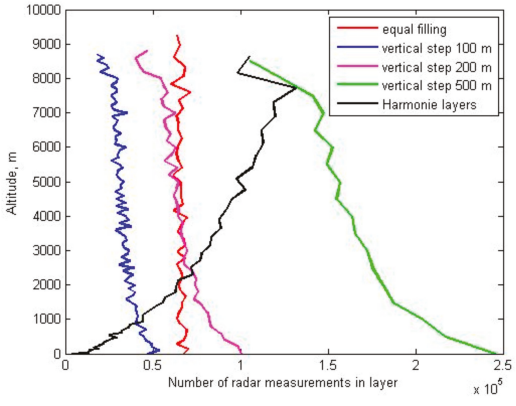


Figure 3. Filling-in the layers by radar data depending on the type of vertical step lists.

2.2. Configuration of Radar Data Assimilation Experiments

Radar reflectivity data from the BaltEx experiment covering the Finnish domain were used in this study (www.baltex-research.eu). The BaltRad project (se.baltrad.eu) related to these data points operates with high-quality data and has demonstrated to possess a significant value for forecasters and decision-makers. The Finnish domain was chosen due to several reasons. First, a major part of the domain is covered by radar measurements even though their inhomogeneous distribution is a factual issue. Second, the smooth orography and relatively homogeneous surface significantly decrease a potential external impact from sharp gradients of the complex orography and surface contrasts.

The experimental design has been implemented on the operational Harmonie platform as follows. A heavy precipitation event over Finland during 14–15 August 2010 with the maximum rain rate during the period 15:00–18:00 UTC was considered. Numerical experiments were performed in the framework of the mesoscale Harmonie-40h1.1 model with 2.5 km domain resolution. Three runs have been carried out with the same model configuration including 3D-Var data assimilation with a three-hour updating cycle. The differences between experiments were in only the radar data involvement and the pre-processing procedure. The control run (CNTR) included the assimilation of all the available SYNOP, TEMP, and AIREP but RADAR observation types over the domain. Two radar reflectivity data assimilation runs used the following approaches. The first approach proposes "cube-smoothing" (FINE) described above with the horizontal resolution equal to the model's 2.5 km resolution and vertical resolution, according to quasi-equivalent filling. The second one (COARSE) explores the

operationally implemented procedure based on double subsequent smoothing of the radar network with resolutions of 8 km and 15 km, respectively, while the model resolution was still kept to 2.5 km [14].

3. Results of Radar Data Assimilation

The results of the numerical experiments have revealed a noticeable impact of radar reflectivity data assimilated in the model on the prediction of precipitation patterns and rates. Moreover, it has been shown that this impact varies depending on pre-processing procedures and their corresponding parameters. Figure 4 shows the spatial distribution of precipitable water and its differences in experiments over the Finnish domain for 18:00 UTC 14 August 2010. This variable was chosen since it better pronounces continuous distribution of water in the atmosphere compared to the discrete form of surface precipitation (Figure 5). Even though both variables reflect major features of spatial coverage, certain discrepancies can be associated with the resolution in radar data assimilation (RDA) formulations. In particular, both methods of RDA lead to redistribution of precipitation toward fine-scale cells within large areas accompanied by an enlarging area of increased precipitation in total. The COARSE run provides larger amounts of precipitation than the FINE run. However, special attention would be paid to the South Finland area where six radars (Anjalankoski, Ikaalinen, Korpo, Kuopio, Vantaa, and Vimpeli (en.ilmatieteenlaitos.fi/fmi-radar-network)) account for the dense three-dimensional coverage of the water content field in the atmosphere. Analysis of the CNTR run with the operationally used configuration shows a rather homogeneous field of precipitable water over this area (see Figure 4a). RDA allows us to specify and redistribute precipitable water among several areas since it is displayed in Figure 4. The first area is associated with the increased precipitation rate in both RDA runs (area I in Figure 4b,c) while the gain is sufficiently higher in the FINE run. This is reflected by the negative values in Figure 4d (COARSE-FINE runs). Over the second area (area II in Figure 4b,c), both RDA runs decrease the water content in the atmosphere, which is actually reallocated toward the first neighboring one. As in the previous site, changes are larger in the FINE run, which result in the area of positive values in Figure 4d. A belt of lower precipitable water amount in the FINE run corresponds to the third area (area III in Figure 4b,c), which, however, does not appear in the COARSE run.

Bearing in mind the above, it can be stated that RDA affects the prognostic fields of the precipitation rate and spatial distribution. In general, this leads to slightly increasing in total precipitation amount over the domain while redistribution of water in the atmosphere occurs in a form of fine-scale cells of opposite signs within a common area of precipitation. As expected, the value and size of these changes depend on the smoothing parameter. The larger this parameter, the smoother the impact in spatial coverage is and the smaller the deviations from the analysis are [21]. Matching the radar resolution to that of the model should minimize numerical effects while focusing on physical features. In particular, the belt and cell features in precipitation fields [22,23] are better simulated with RDA. Higher vertical resolution used in "cube-smoothing" RDA allows us to correct the position of the condensation level, which reduces the phase error. However, the impact from RDA is numerically sensitive to the radar pre-processing approaches and their internal parameters.

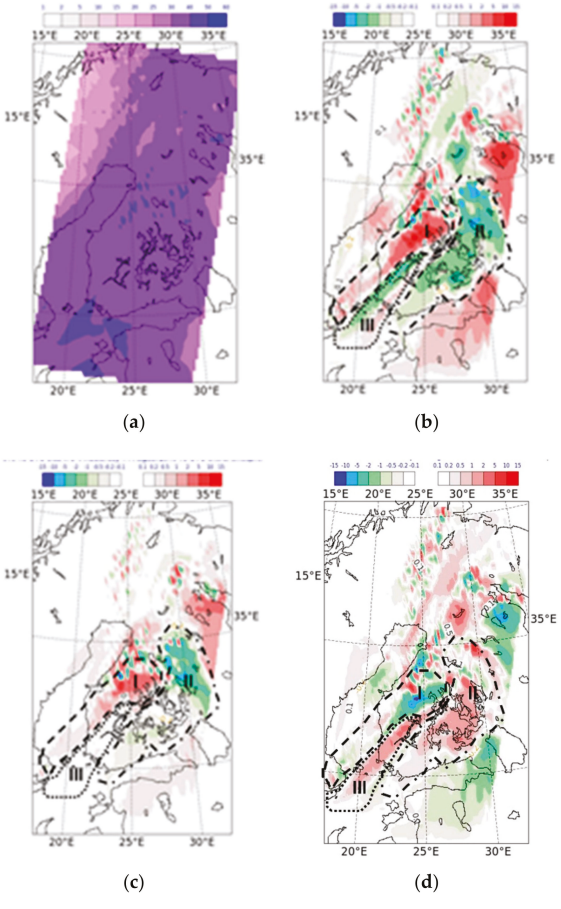


Figure 4. Spatial distribution of precipitable water over the Finnish domain for 18:00 UTC 14 August 2010. (**a**) in analysis and differences between runs: (**b**) FINE-CNTR, (**c**) COARSE-CNTR, (**d**) COARSE-FINE. Areas I, II, and III outline specific regions in redistributing of precipitable water.

The RDA procedures change also the vertical distribution of rain water in the atmosphere. Figure 6 shows the vertical profiles for this variable in the area of the largest differences between the FINE and COARSE runs. The former is similar, in its general features, to the control run, but slightly redistributes water from a 2.5 to a 3.5 km layer towards lower levels. This can produce rare ripe condensation and a decrease of downstream precipitation in favor of upstream, which is shown in Figure 4b. The COARSE run shows a significant reduction of water content in the area, which is explained by the following two reasons. The first is that the deep smoothing spreads a compact substance (water cell) over larger neighboring areas and decreases its concentration. The second, which is related to the above, is that the smoothing changes the water (vapor) concentration, which follows changes in a chain of microphysical processes. The detailed study of links in this chain is a subject of ongoing research.

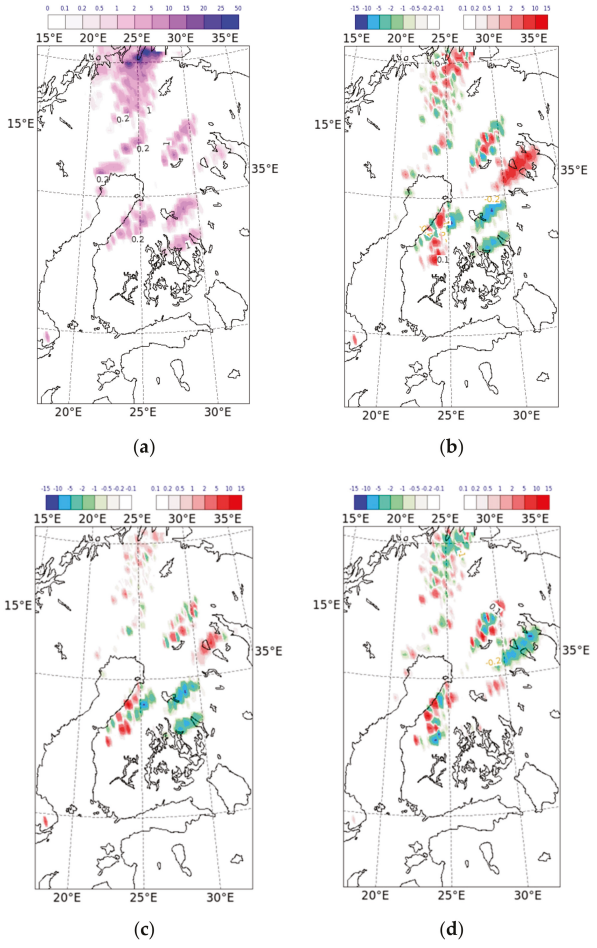


Figure 5. Same as Figure 4, but for surface precipitation.

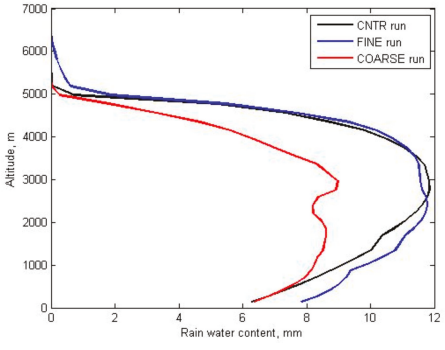


Figure 6. Vertical profiles of rain water in the atmosphere for different runs.

4. Discussion and Concluding Remarks

The quality of numerical weather forecasts is crucially dependent on several factors. The data assimilation remains in the first line among these factors since this determines the error in the initial conditions and the further error growth rate during a forecast. The problem becomes even sharper for mesoscale atmospheric modeling with a resolution of order of a kilometer. Conventional data are less useful because of their coarse spatial and temporal resolutions. Remote sensing is a way for improving this situation due to the fact that it provides high resolution measurements of the order of a few hundred meters over the area of interest with a few minutes of a discretization time-step. However, another sort of problem appears in the forefront: this approach yields an abundance of information, which has to be screened from a useful core. Thus, a compromise has to be made between a volume of detailed information and a restricted amount of selected data, which, however, should still keep spatial and temporal atmospheric features. This is a subject that is dealt with the preprocessing step before data are forwarded to the assimilation module of a numerical atmospheric model.

In this research, the data assimilation system in the convection-permitting Harmonie model has been further developed by involving radar reflectivity measurements. The focus has been on optimizing the inner parameters of the pre-processing procedures. In pursuing the compatibility between the model resolution and smoothed radar observation density, the "cube-smoothing" approach has been proposed, which promotes the presentation of radar measurements on a regular grid at a resolution equal to that of the model's grid. This ensures the equivalent presentation of precipitation (reflectivity) structures in both spaces such as the model and observation in the sense of equally preserving the scales of precipitation patterns. From the spectral point of view for spatial frequencies (wave numbers), this implies an equal cut-off of a high-frequency band in the model and in the observation fields. In this way, at least one component of the forecast error associated with the representativeness error [17] has been minimized.

As shown from the results presented above, the numerical experiments with the proposed approach exhibit better simulation of mesoscale cells and belts of precipitation versus the respective control runs. In particular, instead of one large area with smoothed values for rain rates, the improved assimilation approach was allowed to specify and redistribute precipitable water among several sub-areas. Two of them have higher accumulated water content. The third area located between these has lost some precipitable water. Such redistribution retrieves mesoscale features with sizes from a few to ten kilometers in precipitation fields and improves rain rates over the domain. However, precise verification still remains an issue due to several reasons including being the instrumental error and transient functions for radar measurements as well as the dominating sizes of precipitation patterns in particular regions and under certain atmospheric flow regimes. The latter will be the focus of further studies.

Author Contributions: S.I. contributed to the conceptualization, methodology, formal analysis, investigation, and writing; S.M. contributed to the supervision, methodology, writing, review and editing; I.R. contributed to the formal analysis, investigation and visualization.

Funding: This research received no external funding.

Acknowledgments: The research was realized under support of the Enviro-PEEX on ECMWF project. The authors thank Alexander Mahura (FMI), who is the project principal investigator, for facilitating our involvement in the project and Mats Dahlbom (DMI), who is the author of the preporepapy script, for fruitful discussions.

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

References

1. Gao, J.; Tang, G. Similarities and improvements of GPM Dual-Frequency Precipitation Radar (DPR) upon TRMM precipitation radar (PR) in global precipitation rate estimation, type classification and vertical profiling. *Remote Sens.* **2017**, *9*, 1142. [CrossRef]
2. 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]
3. 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]
4. Bai, L.; Shi, C.; Li, L.; Yang, Y.; Wu, J. Accuracy of CHIRPS satellite-rainfall products over mainland China. *Remote Sens.* **2018**, *10*, 362. [CrossRef]
5. Kidd, C.; Bauer, P.; Turk, J.; Huffman, G.; Joyce, R.; Hsu, K.-L.; Braithwaite, D. Intercomparison of high-resolution precipitation products over Northwest Europe. *J. Hydrometeorol.* **2012**, *13*, 67–83. [CrossRef]
6. Huuskonen, A.; Saltikoff, E.; Holleman, I. The operational weather radar network in Europe. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 897–907. [CrossRef]
7. Ducrocq, V.; Lafore, J.-P.; Redelsperger, J.-L.; Orain, F. Initialization of a fine-scale model for convective-system prediction: A case study. *Q. J. R. Meteorol. Soc.* **2000**, *126*, 3041–3065. [CrossRef]
8. Ducrocq, V.; Ricard, D.; Lafore, J.-P.; Orain, F. Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Weather Forecast.* **2002**, *17*, 1236–1256. [CrossRef]
9. Richard, É.; Cosma, S.; Tabary, P.; Pinty, J.-P.; Hagen, M. High-resolution numerical simulations of the convective system observed in the Lago Maggiore area on the 17 September 1999 (Map IOP 2a). *Q. J. R. Meteorol. Soc.* **2003**, *129*, 543–564. [CrossRef]
10. Lorenc, A.C. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **1986**, *112*, 1177–1194. [CrossRef]
11. Seko, H.; Kawabata, T.; Tsuyuki, T.; Nakamura, H.; Koizumi, K. Impacts of GPS-derived water vapor and radial wind measured by Doppler radar on numerical prediction of precipitation. *J. Meteorol. Soc. Jpn.* **2004**, *82*, 473–489. [CrossRef]
12. Salonen, K.; Järvinen, H.; Haase, G.; Niemelä, S.; Eresmaa, R. Doppler radar radial winds in HIRLAM. Part II: Optimizing the super-observation processing. *Tellus* **2009**, *61A*, 288–295. [CrossRef]
13. Desroziers, G.; Ivanov, S. Diagnosis and adaptive tuning of information error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.* **2001**, *127*, 1433–1452. [CrossRef]
14. Caumont, O.; Ducrocq, V.; Watterlot, É.; Jaubert, G.; Pradier-Vabre, S. 1D+3DVar assimilation of radar reflectivity data: A proof of concept. *Tellus* **2010**, *62A*, 173–187. [CrossRef]
15. Jung, Y.; Xue, M.; Tong, M.J. Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one-and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Mon. Weather Rev.* **2012**, *140*, 1457–1475. [CrossRef]
16. Gao, J.; Xue, M.; Stensrud, D.J. The development of a Hybrid EnKF-3DVAR algorithm for storm-scale data assimilation. *Adv. Meteorol.* **2013**, *2013*, 12. [CrossRef]
17. Ivanov, S.; Palamarchuk, J. Fine-scale precipitation structure of a cold front and the problem of the representativeness error. *Adv. Geosci.* **2007**, *10*, 3–8. [CrossRef]
18. Thepaut, J.-N.; Courtier, P.; Belaud, G.; Lemaitre, G. Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Q. J. R. Meteorol. Soc.* **1996**, *122*, 535–561. [CrossRef]
19. Rabier, F. Overview of global data assimilation developments in numerical weather-prediction centres. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3215–3233. [CrossRef]
20. Ivanov, S.; Simmer, C.; Palamarchuk, J.; Bachner, S. The systematic error of precipitation and humidity in the MM5 model. *Adv. Geosci.* **2008**, *16*, 97–107. [CrossRef]
21. Nogueira, S.M.C.; Moreira, M.A.; Volpato, M.M.L. 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]

22. Tweedy, D.L. Mesoscale precipitation patterns in extratropical cyclones and implications for cyclone development. *J. Geophys. Res. Atmos.* **1990**, *95*, 1987–1997. [CrossRef]
23. Wong, M.; Skamarock, W.C. Spectral characteristics of convective-scale precipitation observations and forecasts. *Mon. Weather Rev.* **2016**, *144*, 4183–4196. [CrossRef]

remote sensing

MDPI

Article

Assessment of Level-3 Gridded Global Precipitation Mission (GPM) Products Over Oceans

Sana Khan * and Viviana Maggioni

Sid and Reva Dewberry Department of Civil, Environment, and Infrastructure Engineering, George Mason University, Fairfax, VA 22030, USA; vmaggion@gmu.edu

* Correspondence: skhan51@gmu.edu

Received: 17 December 2018; Accepted: 24 January 2019; Published: 27 January 2019

Abstract: The performance of Level-3 gridded Global Precipitation Mission (GPM)-based precipitation products (IMERG, Integrated Multi-satellite Retrievals for GPM) is assessed against two references over oceans: the OceanRAIN dataset, derived from oceanic shipboard disdrometers, and a satellite-based radar product (the Level-3 Dual-frequency Precipitation Radar, 3DPRD). Daily IMERG products (early, late, final) and microwave-only (MW) and Infrared-only (IR) precipitation components are evaluated at four different spatial resolutions (0.5°, 1°, 2°, and 3°) during a 3-year study period (March 2014–February 2017). Their performance is assessed based on both categorical and continuous performance metrics, including correlation coefficient, probability of detection, success ratio, bias, and root mean square error (RMSE). A triple collocation analysis (TCA) is also presented to further investigate the performance of these satellite-based products. Overall, the IMERG products show an underestimation with respect to OceanRAIN. Rain events in OceanRAIN are correctly detected by all IMERG products ~80% of the times. IR estimates show relatively large errors and low correlations with OceanRAIN compared to the other products. On the other hand, the MW component performs better than other products in terms of both categorical and continuous statistics. TCA reveals that 3DPRD performs consistently better than OceanRAIN in terms of RMSE and coefficient of determination at all spatial resolutions. This work is part of a larger effort to validate GPM products over nontraditional regions such as oceans.

Keywords: precipitation; satellite remote sensing; error analysis; triple collocation

1. Introduction

Global precipitation data play a key role across numerous applications such as hazard mitigation, terrestrial hydrology, climate change studies, as well as agriculture and irrigation practices [1–4]. The National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) launched the Global Precipitation Measurement (GPM) mission in 2014, which carries an advanced radar (Dual-frequency Precipitation Radar; DPR) and a radiometer system (GPM Microwave Imager, GMI) to measure global precipitation from space [5]. Certain geographical locations, such as oceans and impervious terrain regions, entirely depend on satellite measurements as the only source of precipitation information [6] and the Integrated Multi-satellite Retrievals for GPM (IMERG) products represent the state-of-the-art product for global precipitation estimation. The IMERG algorithm fuses information from several satellite-based microwave and infrared precipitation estimates, as well as ground gauge information.

The measurement and validation of oceanic precipitation are vital for understanding the global water cycle, as most of the global precipitation occurs over oceans [7]. Due to the scarcity of surface measurements over oceans, satellite-based precipitation observations often represent the only source of information. It is critical to characterize errors associated with these products, as these errors may

lead to erroneous conclusions in many applications, e.g., fresh water budgets, currents, cyclones and hurricanes propagation, and El Niño-Southern Oscillation (ENSO) cycle predictions. The GPM Ground Validation (GV) group utilizes reference data obtained from multiple sources, including radar networks and ground-based measurements to validate satellite precipitation measurements [8]. However, validating satellite products over oceans is more challenging because of their inaccessibility and geographical extent (~75% of earth's surface). Past efforts have used weather radars situated on islands and coastlines [9]. Over the years, rain gauge measurements onboard cruise, merchant, and research ships have been used to estimate oceanic precipitation in addition to tropical buoy gauge arrays and manual observations from voluntary observation ships [10–12]. Rain gauges mounted on buoys take point measurements with high temporal resolution, but are easily influenced by wind speed, which can lead to erroneous estimates. In contrast, satellite precipitation estimates have relatively high spatial coverage but lower temporal resolution [13].

Thanks to the launch of the Tropical Rainfall Measurement Mission (TRMM) in 1997 and GPM in 2014, satellite precipitation products have provided useful information on oceanic precipitation quantification and patterns. Bowman [14] and Serra and McPhaden [15,16] compared TRMM-based precipitation retrievals against rain gauge data from ocean buoys in the Pacific and Atlantic Oceans. These studies highlighted the uncertainties associated with an area-averaged and a point measurement-based comparison. In order to fill this gap, scientific groups, such as the International Precipitation Working Group (IPWG) and the GPM GV group, have pressed the need for improved and robust oceanic precipitation surface measurements and associated errors [17]. OceanRAIN (Ocean Rainfall and Ice-phase precipitation measurement Network) is an effort to provide high quality along-track shipboard precipitation data for surface validation over global oceans [18]. IMERG V03 has been evaluated against OceanRAIN precipitation and showed an underestimation of shallow tropical rainfall [19]. Another recent study used OceanRAIN precipitation data for evaluating the HOAPS (Hamburg Ocean Atmosphere Parameters and fluxes from Satellite data) precipitation product across the Atlantic Ocean [20]. This study highlighted that the differences between HOAPS and OceanRAIN are governed by the point-to-area (along track-pixel) effect rather than the precipitation regime itself. Although some efforts have been done in this direction, there is still a need for more evaluation of satellite-based products (particularly their most recent versions) over oceans.

The overall objective of this study is to investigate and better understand how the latest Version 05 of the IMERG products and components performs across oceans at different spatial resolutions. As there is no reference that is continuously available both in time and space over oceans, IMERG products are evaluated against the OceanRAIN precipitation dataset and the Level-3 DPR product, 3DPRD. This work assesses all products at the daily temporal resolution and at four different spatial resolutions, i.e., 0.5°, 1°, 2° and 3°. The GPM GV group lays special emphasis on validation studies in "nontraditional" regions like oceans and this study fits well within this wider effort. The datasets and methodology used in this study are presented in Section 2. Results are described in Section 3 and discussed in Section 4, whereas Section 5 summarizes the main conclusions and presents future research directions.

2. Materials and Methods

2.1. Data

IMERG fuses information from multiple sources, including satellite microwave and infrared precipitation estimates and ground gauge information. The IMERG algorithm uses several passive microwave (PMW) sensors to assemble and intercalibrate precipitation estimates. However, due to the limited sampling of PMW sensors on low-earth-orbit platforms, the gaps are filled by MW-adjusted IR estimates. As IMERG products are provided at 30 min temporal resolution, MW retrievals are obtained either from calibrated conical-scan microwave radiometer or cross-track-scan microwave retrieval, essentially from the one that is closer to the half-hour window. IR estimates are provided

by various sensors (refer to [21] for more details). The IMERG algorithm is run twice in near-real time and once after monthly gauge data become available to generate the following multi-satellite products: (i) early, available ~4 h after the observation time allowing quick assessment for flood and landslide forecasts; (ii) late, available ~12 h after the observation time, mainly for agricultural applications like drought monitoring, crop yield forecasts, and crop production; and (iii) final, available ~2 months after observation for research applications. The data fields are provided by the NASA Precipitation Processing System (PPS) as global Level-3 gridded precipitation estimates ($mm \cdot h^{-1}$) at the spatial/temporal resolution of 0.1°/30 min. In this study, the following data fields archived in each IMERG file are used: *precipcal* for the early, late, and final merged products, *HQprecipitation* for the MW component, and *IRprecipitation* for the IR component.

The GPM core satellite carries the dual-frequency precipitation radar operating at Ku-band (13.6 GHz) and Ka-band (35.5 GHz). DPR provides measurements that are spaced at ~5 km, over 245 km (at Ku-band) and 120 km (at Ka-band) wide swaths. The attenuation and loss of the radar signal along with precipitation phase still influence DPR observations, as in the TRMM Precipitation Radar [22–25], but these issues are addressed in the Level-2 DPR algorithm. The Level-3 DPR algorithm performs a spatial and temporal statistical processing to generate gridded products that cover the whole globe and are available at multiple temporal (hourly, daily and monthly) and spatial (0.25° and 0.50°) resolutions. In this work, the daily Level-3 product 3DPRD is used at its native spatial resolution of 0.25° and with a spatial coverage from 67° S to 67° N and from 180° W to 180° E. The data archived in the *GRID/precipRateESurfMean* data field are utilized in this study. IMERG uses Level-2 PMW precipitation retrievals trained by the 2BCMB (GMI+DPR) combined algorithm. For interpreting reflectivity profiles from DPR reflectivity observations, 2BCMB uses a different approach than the DPR algorithm to interpret the reflectivity profiles. Moreover, several levels of processing from the 2BCMB estimates to the Level-2 PMW, which adopts a Bayesian retrieval algorithm, to IMERG guarantees that the IMERG product is different from the 2BCMB precipitation on a scene-by-scene basis, although their long-term means will look similar as a result of the intercalibration [26].

Several research vessels (RV) have gathered precipitation data along with auxiliary atmospheric data as part of the OceanRAIN initiative over the Pacific and Atlantic Oceans since 2010. The mainstay of OceanRAIN is an optical disdrometer (the Eigenbrodt ODM470) coupled with an anemometer and a precipitation detector. The measuring volume is illuminated evenly along the cross section by an infrared light diode. The ODM470 records the light extinction caused by the passing hydrometeor in terms of activation voltage which is proportional to the cross-sectional area of the hydrometeor [27]. To obtain the particle size distribution, the hydrometeors are counted in a 60 s window and sorted in a bin size ranging from 0.04 mm to 22 mm. A particle size distribution is then used to derive precipitation phase, intensity, accumulation, and precipitation occurrence. Moreover, the installation height of ODM470 on RVs assures minimal splashing water effect on the measurements. Further details regarding the instrument and the OceanRAIN dataset can be found in Klepp et al. [28]. Figure 1 shows the ship tracks of the RVs collecting precipitation data during the 3-year study period (March 2014–February 2017).

Information regarding the sample size of precipitation data collected by each RV is presented in Table 1. The temporal resolution of the data is 1 min. The ODM470 instrument used in OceanRAIN can measure very light precipitation (from 0.01 $mm \cdot h^{-1}$ to 0.09 $mm \cdot h^{-1}$) unlike common gauges. As highlighted in Klepp et al. [29], when an insignificant number of particles is recorded, the measurement is flagged and assigned a value of 0.00 $mm \cdot h^{-1}$. Moreover, to avoid unrealistic spikes, quality control is maintained through routine visual inspections. Seven RVs have acquired OceanRAIN data during the three years of analysis. Polastern has the longest time series and thus collected the largest numbers of samples, whereas World spans over two months only and shows the smallest sample size. Measurements from all RVs are merged together to facilitate the comparison with the IMERG products/components, resulting in a total of 623,641 samples over the 3-year period. Considering rain rates larger than 0 $mm \cdot h^{-1}$, the sample size reduces to 379,744 during the same study

period. IMERG, DPR, and OceanRAIN data are all matched to the daily temporal scale and to a regular cartesian grid at four different spatial resolutions (0.5°, 1°, 2° and 3°).

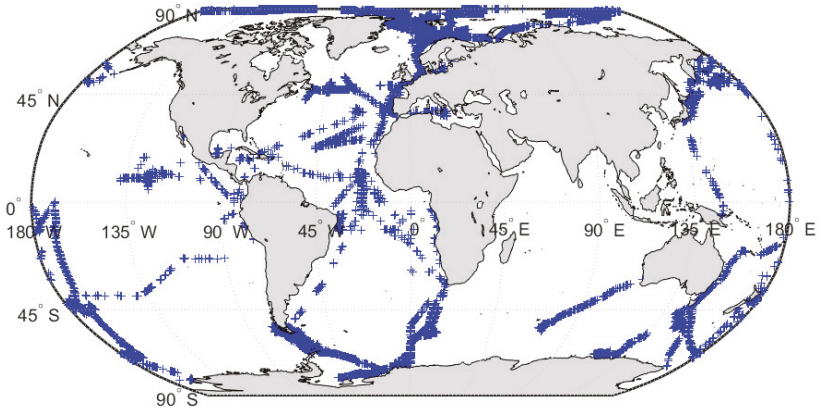


Figure 1. Ship tracks of OceanRAIN across the Pacific and Atlantic Oceans during March 2014–February 2017 (blue crosses correspond to 8030 unique locations).

Table 1. Data collected by each research vessel (RV) during the 3-year study period (March 2014–Feb 2017).

RV ID	Sample Size	# Cases with Rain > 0 mm·h^{-1}	Time Series
Sonne-II	60,168	58,083	Nov 2014–Jan 2015
Roger	10,763	8469	Aug–Sep 2014
Polastern	445,635	233,349	Mar 2014–Oct 2016
Meteor	20,299	19,358	Mar 201–Mar 2016
Maria	27,083	23,954	Mar–Jun 2014
Investigator	54,814	32,671	Jan 2016–Feb 2017
World	4879	3860	Jan–Feb 2017
TOTAL	623,641	379,744	Mar 2014–Feb 2017

2.2. Spatio-Temporal Data Alignment

IMERG (both merged products and components), DPR, and OceanRAIN data have different native resolutions. Moreover, OceanRAIN data are not gridded, which also necessitates a spatial alignment to a common grid. IMERG (native resolution = 0.1°) and 3DPRD (native resolution = 0.25°) are mapped to a common 0.5° cartesian grid through simple spatial averaging. Then, all OceanRAIN measurements falling within each 0.5° grid box are averaged together. Both IMERG and OceanRAIN datasets are averaged temporally for every day during the 3-year study period to match the daily resolution of 3DPRD.

In order to investigate the impact of spatial resolution (i.e., grid box size) on our analysis, daily IMERG and 3DPRD data are upscaled to coarser resolutions (1°, 2°, and 3°) and compared to the OceanRAIN data. IMERG, 3DPRD, and OceanRAIN at these resolutions are generated by spatially averaging all the valid observations (including zeros) available within each grid box. Statistical metrics for the three collocated products are thus computed at a total of four spatial resolutions (0.5°, 1°, 2° and 3°) and at the daily time scale.

2.3. Performance Analysis

Categorical and continuous statistics allow characterization of systematic and random errors, which are both critical to evaluate and further improve precipitation retrieval algorithms, for instance through bias adjustment techniques. Both categorical and continuous statistics are used to assess the performance of the IMERG products/components against two references, i.e., OceanRAIN and 3DPRD. The rationale behind the use of two different reference datasets consists in the fact that over oceans there is no ideal benchmark for evaluating precipitation products that is continuously available both in time and space. A rain/no-rain threshold of 0.01 mm·h^{-1} is used to compute categorical and continuous statistics. Categorical statistics include probability of detection (POD), success ratio (SR), critical success index (CSI), and the hit bias. Among the continuous statistics, correlation coefficient (CC), standard deviation (SD), and the root mean-square error (RMSE) are considered. Definitions and equations for each metric are reported in Appendix A, together with the contingency table on which the categorical statistics are based. All error metrics are computed at the daily scale for the four spatial resolutions.

Triple collocation is an alternative method for assessing the quality of a product without assuming a reference and the random errors are computed against an unknown truth [30–36]. This technique has been widely used in precipitation studies, although some studies showed that TCA is highly sensitive to its input configurations, including scale differences, time span under observation, and measurement triplets [37–39]. The method proposed by McColl et al. [31] is applied here to estimate RMSE and R^2 values of a specific precipitation product with the truth defined as:

$$X_i = X_i' + \varepsilon_i = \alpha_i + \beta_i t + \varepsilon_i \tag{1}$$

where $X_i (i \in \{1, 2, 3\})$ represents collocated measurement systems linearly related to the true value t with additive random errors ε_i, whereas α_i and β_i represent the least square intercepts and slopes, respectively. Assuming that the errors from each system have (i) zero mean, (ii) are mutually uncorrelated ($\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j$), and (iii) orthogonal with respect to t ($\mathrm{Cov}(\varepsilon_i, t)$), the RMSE and coefficient of determination (R^2) are computed as follows:

$$RMSE = \begin{bmatrix} \sqrt{Q_{11} - \frac{Q_{12}Q_{13}}{Q_{23}}} \\ \sqrt{Q_{22} - \frac{Q_{12}Q_{23}}{Q_{13}}} \\ \sqrt{Q_{33} - \frac{Q_{12}Q_{23}}{Q_{12}}} \end{bmatrix} \tag{2}$$

$$R^2 = \begin{bmatrix} \frac{Q_{12}Q_{13}}{Q_{11}Q_{23}} \\ \frac{Q_{12}Q_{23}}{Q_{13}Q_{22}} \\ \frac{Q_{13}Q_{23}}{Q_{12}Q_{33}} \end{bmatrix} \tag{3}$$

where $Q_{ij} = \mathrm{Cov}(X_i, X_j)$ and Cov represents the covariance matrix.

We applied TCA by aggregating daily data at different spatial resolutions (0.5°, 1°, 2° and 3°) over the 3-year period and by selecting instances when and where all three datasets have precipitation rates $\geq$ 0.01 mm·h^{-1}. TCA is applied to the following triplets of precipitation products: (1) early-3DPRD-OceanRAIN (Triplet A); (2) late-3DPRD-OceanRAIN (Triplet B); (3) final-3DPRD-OceanRAIN (Triplet C); (4) MW-3DPRD-OceanRAIN (Triplet D); and (5) IR-3DPRD-OceanRAIN (Triplet E). R^2 and the RMSE are then computed for each triplet.

3. Results

As a first step to investigate the datasets considered in this study, the probability density functions (PDFs) of each of the six daily precipitation products at 0.5°, 1°, 2° and 3° spatial resolutions are investigated (Figure 2). At the finer scale (0.5°), 3DPRD demonstrates a sharp distribution concentrated

around small precipitation rates (<0.01 mm·h^{-1}), whereas IR reveals the most uniform distribution among all datasets. The MW component and the merged products have all similar distributions that are close to the one of OceanRAIN. Coarser resolutions result in flatter 3DPRD distributions that get closer to the other IMERG products. In the IR distributions, the density of low precipitation rates becomes larger at coarser resolutions. Other IMERG products maintain similar distributions when moving from finer to coarser resolutions, except for IMERG early, which flattens out significantly at 3°. The distribution of OceanRAIN is similar across the four resolutions, albeit slight variations at small rain rates (<0.01 mm·h^{-1}).

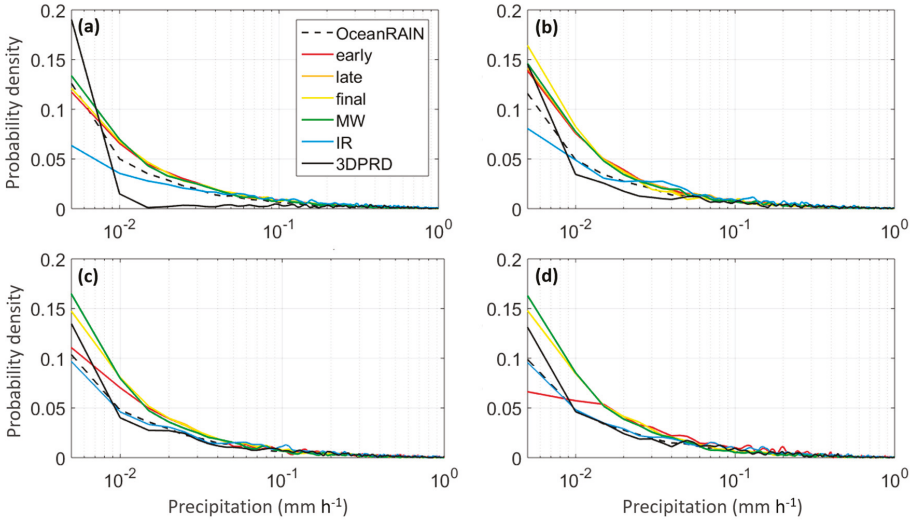


Figure 2. Probability density functions of GPM precipitation products and OceanRAIN precipitation over the Pacific and Atlantic Oceans during March 2014–February 2017 at (**a**) 0.5°, (**b**) 1°, (**c**) 2°, and (**d**) 3° spatial resolution.

In order to further investigate the performance of the satellite-based precipitation products with respect to the ground reference, cumulative distribution functions (CDFs) of daily IMERG products and 3DPRD are compared to the CDF of OceanRAIN at 0.5° and 3° resolutions (Figure 3). All IMERG products and components show an overall underestimation with respect to OceanRAIN, both at finer and coarser spatial resolutions. This underestimation increases when moving from the finer to the coarser resolution. When analyzing the CDF of 3DPRD at 0.5°, we observe an overestimation of precipitation rates ranging between 0.01 mm·h^{-1} and 0.9 mm·h^{-1} and an underestimation at larger rain rates. However, at the coarser resolution, 3DPRD consistently underestimates the reference, similarly to the other satellite-based products.

As mentioned above, the performance of the IMERG products and components is further investigated against two different references: OceanRAIN and 3DPRD. The performance metrics considered in this work are illustrated in Figures 4 and 5 in terms of performance and Taylor diagrams [40,41]. The performance diagram displays categorical statistics such as probability of detection (POD), success ratio (SR), critical success index (CSI), and hit bias. The upper right corner of the diagram represents the perfect score with POD, SR, CSI and hit bias approaching 1. Both performance and Taylor's diagrams are presented at the four different resolutions: 0.5° (a–b); 1° (c–d); 2° (e–f); and 3° (g–h) for each one of the five IMERG precipitation products. The left panels in each Figure present results for OceanRAIN as reference, whereas the right panels show 3DPRD as the reference.

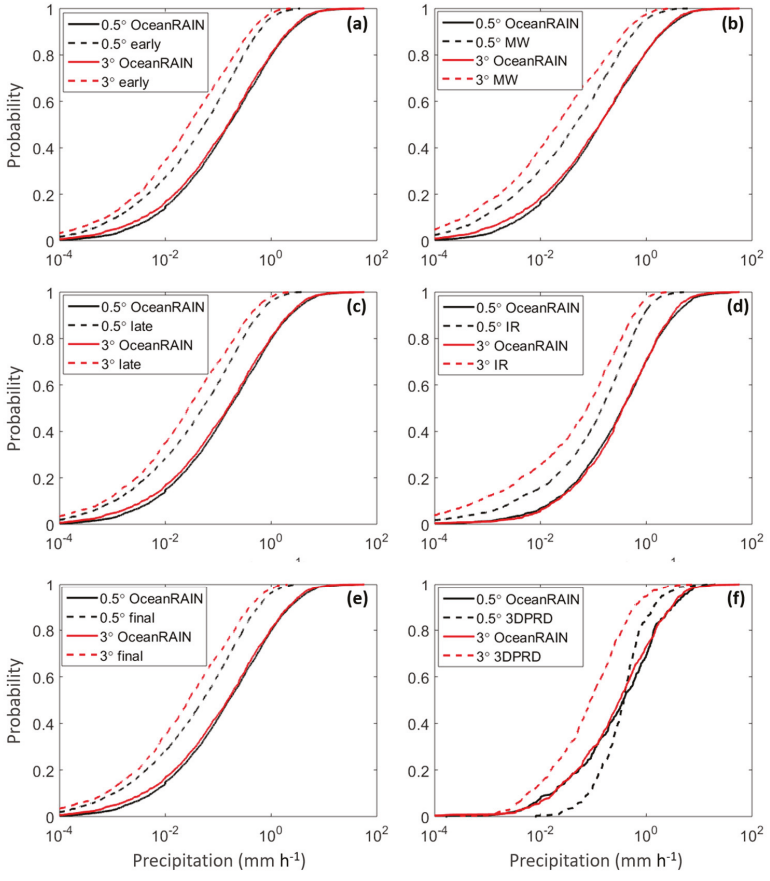


Figure 3. Cumulative distribution functions of OceanRAIN and (**a**) IMERG early; (**b**) MW; (**c**) IMERG late; (**d**) IR; (**e**) IMERG final and (**f**) 3DPRD) at 0.5° (black) and 3° (red) spatial resolution during the study period. Precipitation rates on the x-axis are shown in logarithmic scale. Dashed lines correspond to satellite products, whereas solid lines correspond to OceanRAIN.

By analyzing the performance diagrams in Figure 4, we observe that, overall, IR is the only product whose statistics are markedly different from the other products, which are closely packed together in the diagrams. POD is just below 0.80 for all IMERG products except for IR, which has a POD slightly greater than 0.80 with OceanRAIN as reference. This suggests that all products have a good ability of detecting precipitation (i.e., ~80% of the times that OceanRAIN detects rain, the satellite product agrees). SR on the other hand is the lowest for IR (~0.55) and higher for other products (~0.65). Likewise, the hit bias and CSI are respectively around 1.50 and 0.50 for IR and 1.30 and 0.55 for the other products at 0.5° resolution, when compared to OceanRAIN. The performance metrics show minimal changes with spatial resolutions: SR and CSI exhibit slight improvements at coarser resolutions, whereas POD and hit bias do not show any dependence on the spatial scale.

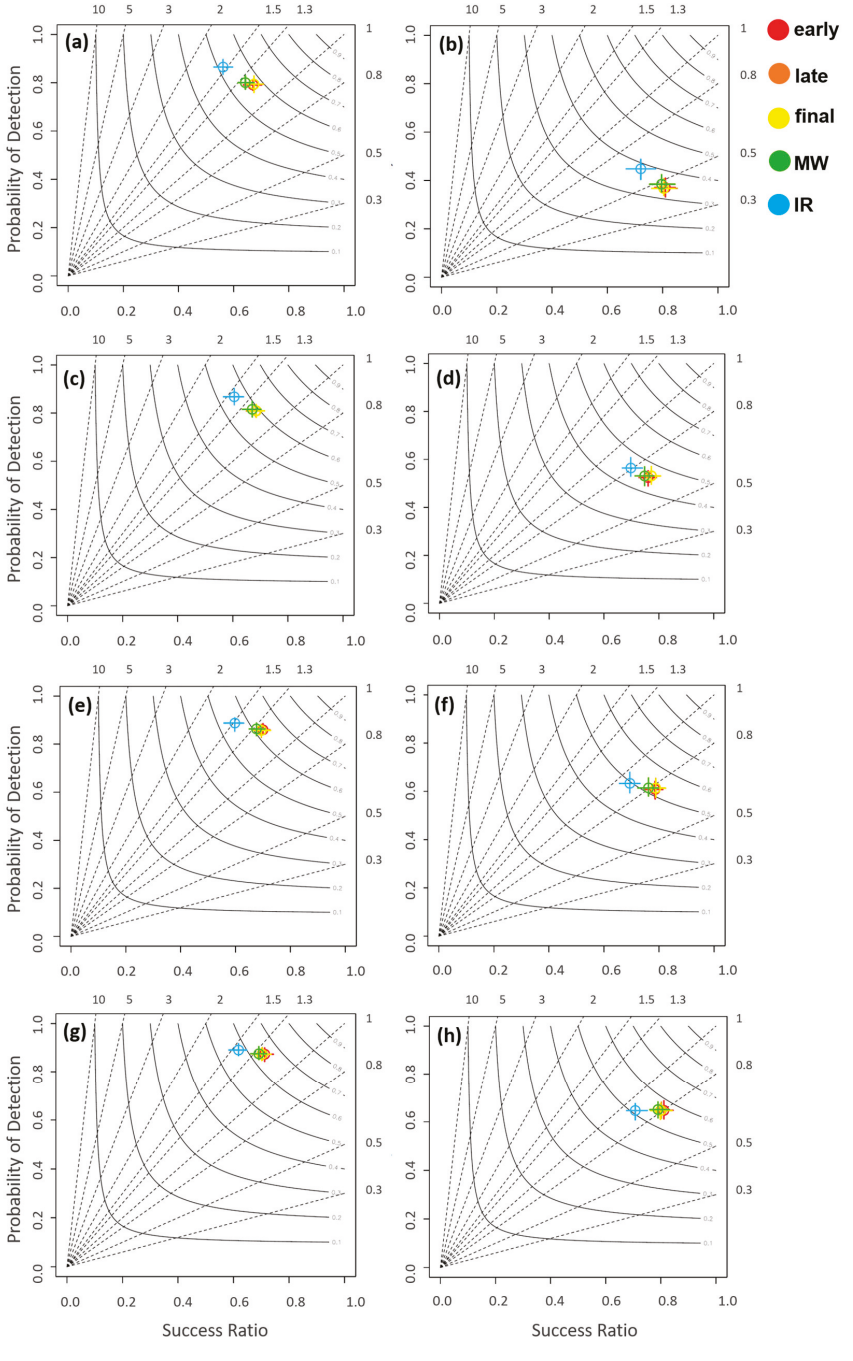


Figure 4. Performance diagrams for the IMERG products vs. OceanRAIN (**a**,**c**,**e**,**g**) and vs. 3DPRD (**b**,**d**,**f**,**h**) at 0.5° (**a**,**b**), 1° (**c**,**d**), 2° (**e**,**f**), and 3° (**g**,**h**) spatial resolutions. Circles represent Probability of Detection (POD) and Success Ratio (SR) for different regions, '+' indicates variance, dotted lines correspond to hit bias, and solid curves to Critical Success Index (CSI) values.

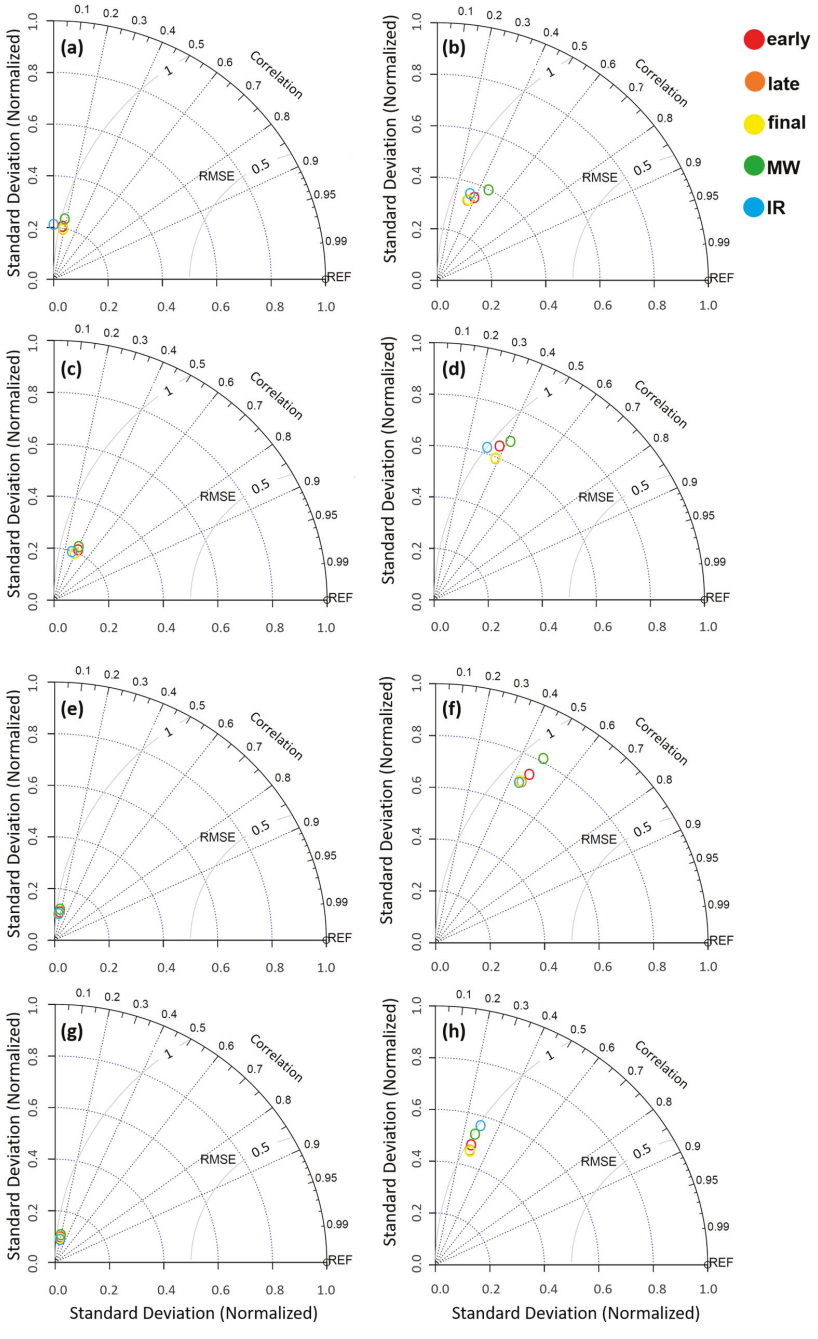


Figure 5. Taylor diagrams for the IMERG products vs. OceanRAIN (**a**,**c**,**e**,**g**) and vs. 3DPRD (**b**,**d**,**f**,**h**) at 0.5° (**a**,**b**), 1° (**c**,**d**), 2° (**e**,**f**), and 3° (**g**,**h**) spatial resolutions. Standard deviation (SD dotted blue curves), RMSE (solid gray curves), and correlation (CC radial dotted black lines) are normalized with respect to the reference data. REF indicates the OceanRAIN/3DPRD-based reference metrics (with itself), i.e., SD and CC of 1 and RMSE of 0.

By analyzing the performance diagrams with 3DPRD as reference (presented in the right panels in Figure 4), we observe that the POD is much lower (0.35–0.65) than in the previous case (when OceanRAIN was taken as benchmark) and that it shows larger variability across different spatial resolutions. However, the SR is not influenced by changes in the spatial resolution and ranges between 0.70 and 0.80 for all products. CSI and the hit bias also manifest improvement at coarser spatial resolutions with values increasing from 0.35 to 0.55 for CSI and from 0.50 to 0.85 for the hit bias. CSI may not always be a reliable indicator because of its dependence on frequency of events. Specifically, it may increase with resolution because of its inherent bias with event frequency. However, CSI is a valid indicator of the relative performance of various algorithms [42]. Interestingly, the difference in performance between IR and other IMERG products follow a very similar pattern for both references.

The Taylor diagrams in Figure 5 present continuous statistics such as correlation (CC), normalized standard deviation (SD), and root-mean square error (RMSE), where RMSE is expressed as a function of CC and SD. A normalized version of the Taylor diagrams, in which the SD and RMSE are normalized by the standard deviation of the reference, is considered here to provide a better measure of the relative performances of the satellite-based precipitation products [41]. The perfect scores are 0 for RMSE and 1 for CC and SD. Similarly to the categorical scores, continuous statistics are computed at four different resolutions: 0.5° (a–b); 1° (c–d); 2° (e–f); and 3° (g–h) spatial resolutions. The left panels correspond to OceanRAIN taken as the reference, whereas the right panels consider 3DPRD as the reference. At 0.5°, the correlation coefficient between all the IMERG products and OceanRAIN is 0.20, except for the IR component, which exhibits minimal CC (0.05). CC doubles for all the products at 1° spatial resolution (5c), including IR, depicting a CC of 0.4. RMSE follows a very similar trend as CC with changing resolutions. Specifically, RMSE shows the lowest values at 1° (0.80), whereas it is 1 for the other resolutions. At 0.5° and 1° SD is 0.25 for all IMERG products and decreases to 0.1 at 2° and 3° resolutions. Overall, the continuous statistics for all IMERG products appear to converge at a single value at 2° and 3°.

Continuous statistics with 3DPRD used as a reference exhibit more pronounced variability at coarser spatial scales. CC changes from 0.40 at 0.5° and 1° resolutions to 0.45 at 2°, and to 0.30 at 3° resolution. RMSE shows values of 0.80 at 0.5° and 2° resolutions and of 1 at 1° and 3° resolutions for all IMERG products. SD moves closer to the reference value (SD = 1) as the spatial resolution goes up from 0.35 (at 0.5°) to 0.75 (at 2°), but it falls back to 0.50 at 3°. Among all IMERG products, the MW estimates exhibit the best continuous statistics when evaluated against 3DPRD.

Since the categorical and continuous statistics do not provide a definitive answer to what reference should be used to evaluate the IMERG products and components, a triple collocation analysis is adopted here as an alternative method to investigate random errors associated with the products of interest. Table 2 summarizes the TCA results for the five triplets for each one of the IMERG products and components, when 3DPRD and OceanRAIN are used as the other two measurements. The RMSE and R^2 values obtained from TCA are compared at the four spatial resolutions. Among the three precipitation measurements in each triplet, the IMERG products consistently exhibit the higher correlations and lower RMSE values.

Figure 6a presents mean values of all precipitation products at four spatial resolutions along the study time series, whereas Figure 6b shows the overall bias (mm·h^{-1}) of each satellite product, defined as follows:

$$Bias = \mu_{Sat} - \mu_{Ref} \quad (4)$$

where μ_{Ref} is the mean precipitation of either OceanRAIN (top panel) or 3DPRD (bottom panel), and μ_{Sat} represents the mean precipitation of each IMERG product/component. Precipitation mean values and biases are computed for collocated products using a threshold of 0.01 mm·h^{-1}.

Table 2. RMSE (in mm·h^{-1}) and R^2 from TCA for five precipitation product triplets at four spatial resolutions (0.5°, 1°, 2°, 3°).

Triplet ID	Products	RMSE (mm·h^{-1}) 0.5°, 1°, 2°, 3°	R^2 0.5°, 1°, 2°, 3°
A	early	0.35, 0.30, 0.22, 0.21	0.50, 0.45, 0.67, 0.53
	3DPRD	1.05, 0.50, 0.40, 0.53	0.32, 0.37, 0.36, 0.19
	OceanRAIN	2.15, 1.61, 3.20, 2.88	0.07, 0.38, 0.05, 0.08
B	late	0.33, 0.29, 0.22, 0.21	0.48, 0.43, 0.62, 0.50
	3DPRD	1.07, 0.49, 0.40, 0.53	0.27, 0.40, 0.38, 0.20
	OceanRAIN	2.14, 1.64, 3.18, 2.88	0.08, 0.35, 0.05, 0.08
C	final	0.32, 0.28, 0.23, 0.20	0.50, 0.43, 0.61,0.51
	3DPRD	1.09, 0.49, 0.40, 0.52	0.26, 0.40, 0.37, 0.21
	OceanRAIN	2.12, 1.62, 3.22, 2.90	0.08, 0.35, 0.05, 0.07
D	MW	0.36, 0.32, 0.24, 0.22	0.62, 0.47, 0.71, 0.53
	3DPRD	1.02, 0.49, 0.40, 0.53	0.37, 0.42, 0.37, 0.19
	OceanRAIN	2.21, 1.68, 3.32, 2.92	0.06, 0.33, 0.05, 0.08
E	IR	0.59, 0.38, 0.29, 0.23	0.04, 0.31, 0.63, 0.48
	3DPRD	1.86, 0.61, 0.48, 0.61	0.01, 0.31, 0.31, 0.12
	OceanRAIN	2.62, 1.87, 3.84, 3.26	0.00, 0.37, 0.04, 0.09

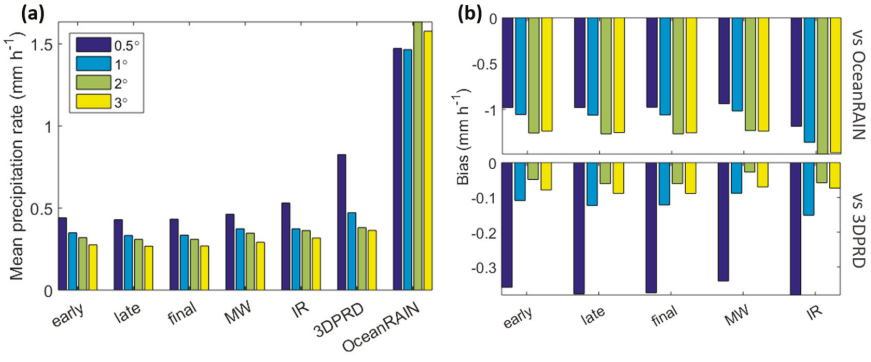


Figure 6. (**a**) Mean precipitation rate (mm·h^{-1}) for all precipitation products and (**b**) bias of the IMERG products against OceanRAIN (top panel) and 3DPRD (bottom panel) during the study period.

The precipitation mean of all satellite products decreases when moving to coarser resolutions (from 0.5° to 3°), likely due to the inclusion of more pixels where no precipitation occurs. However, this is not the case for OceanRAIN, since not as many more observations as in the case of the gridded products are included when increasing resolution. In terms of bias, all IMERG products show larger bias when OceanRAIN is used as a reference as opposed to 3DPRD, which is expected due to the inherent different nature of satellite-based retrievals and track-based observations. For the same reason, biases between the IMERG products and OceanRAIN tend to increase when resolution gets coarser, whereas IMERG biases with respect to 3DPRD decrease with coarser resolution because of the smoothing effect of the spatial aggregation. This is a critical point that should be considered when choosing a satellite-based reference versus a track-based one for evaluating IMERG products over oceans.

TCA also allows us to compare the two references used in the previous analyses against an unknown truth. 3DPRD consistently outperforms OceanRAIN in terms of both R^2 and RMSE at all spatial resolutions. A comparison across the spatial resolutions reveals a very interesting pattern, that is, R^2 values are consistently higher for 3DPRD, as already anticipated by the Taylor diagrams.

In terms of RMSE, mean values for OceanRAIN across all the triplets are 2.25 mm·h^{-1}, 1.68 mm·h^{-1}, 3.35 mm·h^{-1}, and 2.97 mm·h^{-1} at 0.50°, 1°, 2° and 3° resolution, respectively. Mean RMSE value for 3DPRD across all the triplets are 1.22 mm·h^{-1} (0.50°), 0.52 mm·h^{-1} (1°), 2.08 mm·h^{-1} (2°) and 0.54 mm·h^{-1} (3°).

In terms of triplets that consider IMERG products and components, Triplets A-D show comparable performances in terms of both RMSE and R^2, with Triplet D (the one including the MW component) showing the lowest RMSE and the largest R^2 and Triplet E (the one including the IR component) showing the worst performance at all spatial resolutions. It is worth noting that the TCA assumption of linear dependence among the datasets is not met for the 0.50° resolution for the IR component, which leads to R^2 values close to 0 for all the measurements in Triplet E.

4. Discussion

More than 75% of the global rain occurs over oceans, thus necessitating reliable measurements over oceans for better understanding global atmospheric and hydrological processes. The launch of precipitation-focused satellite missions (i.e., TRMM and GPM) in the past few decades has provided the opportunity to estimate precipitation at fine spatial/temporal resolutions over oceans, where surface observations are sparse and often influenced by wind speed and direction and by the stability of the measuring platform [8]. Nevertheless, the true capability of these satellite precipitation estimates over oceans is closely linked to the quantification of the uncertainties associated with them.

In this context, this study assesses the performance of the latest V05 IMERG products and components over oceans. Two independent references (OceanRAIN and 3DPRD) are used to evaluate the products at four spatial resolutions and daily temporal scale over a 3-year study period. All IMERG products show comparable precipitation distributions, except for the IR estimates. An overall underestimation by all IMERG products and 3DPRD against the OceanRAIN is observed for the analysis period, which corroborates results from other studies where moored buoys were used as a reference to evaluate satellite rainfall estimates [13]. Precipitation rates lower than 0.1 mm·h^{-1} dominate the distributions in all products and are more pronounced in 3DPRD estimates. The OceanRAIN PDF appears to migrate closer to the IR PDF as we move to coarser resolutions, which can be attributed to the IR sampling. Specifically, IR has good sampling and despite the noise, the IR mean should converge to the mean of the MW with sufficient spatial/temporal averaging, which could be the case here. The reader should note that the IR field provided in the half hour IMERG file may not have been the one used in the merger with that half hour's satellite estimates. In IMERG versions 03 through 05, the next half hour's IR is used to merge with the current half hour's satellite estimates since it tends to exhibit the highest correlation.

The performance of IMERG products and components is assessed by using a plethora of statistical metrics to provide a comprehensive analysis of their associated uncertainty and errors. Among all products, IR estimates perform worst in terms of CC, SD, and RMSE. However, in terms of POD, the IR component tends to depict slightly improved values as compared to other IMERG products against both references. The IMERG products follow IR in terms of rain detection because IR is the only high-frequency observation contributing to the IMERG algorithm, whereas the MW component is obtained from either calibrated conical-scan microwave radiometer retrievals or cross-track-scan microwave sounder retrievals, whichever is closer to the half-hour window. The IR-based product is estimated indirectly via cloud top temperature measurements. Oceans are commonly characterized by dense cloud cover over tropical and temperate regions [43]. The increased POD for IR estimates could be attributed to the fact that the data collected by RVs are mostly in the tropical and temperate regions typically receiving more convective rainfall, which is well detected by IR sensors [44]. On the contrary, the MW component performs better than the other components in terms of both categorical and continuous statistics, which is in line with past studies conducted using buoys as a reference [13,45]. As this study does not distinguish the Pacific Ocean from the Atlantic Ocean, the performance over a

specific region cannot be inferred. However, previous studies suggest inconsistencies in POD, SR, CC, and RMSE for satellite estimates across the western and eastern Pacific and Atlantic Oceans [45–47].

3DPRD has the potential to be a very useful reference because of its global availability at all times as opposed to OceanRAIN that provides very limited spatio-temporal measurements. In general, while performance metrics are slightly better when OceanRAIN is used as reference to evaluate the IMERG products, the continuous statistics are clearly better when 3DPRD is adopted. This may be attributed to the similar nature of satellite-based products and to the pixel-to-pixel comparison (with respect to the point-to-pixel comparison used in the OceanRain analysis). Although the along-track averaged precipitation of OceanRAIN provides better representation of the areal extent of satellite measurements, the track-to-area difference could be minimized by statistical adjustments as proposed by Burdanowitz et al. [20]. Overall, this does not point to a conclusive evidence as to which reference (OceanRAIN or 3DPRD) should be used for validating the IMERG products. Thus, using an unknown truth-based analysis, such as TCA, is the recommended approach to evaluate random errors associated with satellite-based products over oceans. Theoretically, TCA provides error variance without assuming one of the observations as ground truth. The main assumption of TCA is that three products have mutually independent errors. The results are biased to some extent when two remotely sensed observations are used in the triplets. TCA shows a similar picture in terms of the relative performance of the IMERG products and components with respect to each other as observed in other performance metrics (e.g., RMSE, SD, and CC). Among all products analyzed here, IMERG final shows the lowest RMSE with respect to the truth, followed by late, early, MW, IR, 3DPRD, and OceanRAIN. The R^2 is relatively higher for the IMERG products/components than the one of 3DPRD and OceanRAIN. In TCA, the performance of OceanRAIN is better at 1° resolution compared to other resolutions. However, the satellite-based products do not show this behavior. This can possibly be attributed to the influence of track-to-area averaging, i.e., the areal nature of satellite products vs the track structure of OceanRAIN. This is an interesting point that should be investigated further in future studies.

TCA clearly confirms that 3DPRD performs consistently better than OceanRAIN, as also shown by the continuous error metrics. Khan et al. [48] demonstrated that IR estimates are affected by the largest systematic and random errors over land compared to the other IMERG products and components and their improvement could be critical to enhance the merged products. The findings over ocean also point to larger errors in IR estimates over oceans, thus providing an opportunity to improve the merged products also over oceans.

Using OceanRAIN as a reference for evaluating GPM-based products has its own limitations. First, although OceanRAIN data are available since 2010, only a few years overlap with the GPM mission. Second, measurements are spatially sparse and sample sizes are small at coarser temporal resolutions. At this point mapping errors spatially is not possible because of the short length of the time series, but additional validation analysis should be performed when longer time series of OceanRAIN and GPM estimates will be available. A rain/no-rain threshold value is set to 0.01 $mm \cdot h^{-1}$ for computing categorical and continuous statistics. This threshold may not be ideal for evaluating satellite products at their native resolutions, as it is below the minimum detection limit of most spaceborne instruments. However, a higher threshold would result in a reduced sample size, consequently resulting in less robust conclusions. Future work should investigate the impact of this threshold on the validation analysis.

5. Conclusions

This study evaluates the performance of Level-3 GPM gridded products over oceans using 3DPRD and OceanRAIN as potential references. Similarly, this methodology could be applied to other new generation satellite-based precipitation products, such as the Global Satellite Mapping of Precipitation (GSMaP) product by JAXA [49]. The quasi-global availability of DPR-based measurements makes it a unique reference where and when a surface-based reference is not available. Moreover, compared to

passive remote sensors such as passive microwave and infrared sensors, DPR offers more accurate precipitation estimates. IR estimates exhibit margin for improvement over oceans, which in turn could lead to reduction in errors in the merged products. Future extension of this work could include alternate independent precipitation datasets such as buoys and model reanalysis for the TCA. Future studies should investigate error sources and attempt to separate error components (systematic and random), which is extremely important to implement improvements in satellite precipitation algorithms.

Author Contributions: Conceptualization, S.K. and V.M.; methodology, S.K. and V.M.; data processing, data analysis and visualization, S.K.; writing—original draft preparation, S.K.; writing—review and editing, S.K. and V.M.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank Christian Klepp for providing the OceanRAIN data and NASA Precipitation Processing System (PPS) for making the IMERG and DPR data available for research.

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

Appendix A

Table A1. Contingency table.

		Satellite	
		$P_{Sat} \geq th$	$P_{sat} < th$
Reference	$P_{Ref} \geq th$	H	M
	$P_{Ref} < th$	F	Z

where H represents 'hit' cases, i.e., both the satellite (P_{Sat}) and the reference (P_{Ref}) are greater than or equal to the rain/no-rain threshold (*th*); F represents 'false alarms', i.e., P_{Sat} is greater than or equal to *th*, but P_{Ref} is less than *th*; M represents 'misses', i.e., P_{Ref} is greater than or equal to *th* but P_{Sat} is less than *th*; Z represents 'true negative', i.e., P_{Sat} and P_{Ref} are both less than *th*. The contingency table parameters H, M, F, and Z are defined in Table 1.

$$POD: \ \mathrm{H}/(\mathrm{H}+\mathrm{M}) \tag{A1a}$$

$$SR: \ \mathrm{H}/(\mathrm{H}+\mathrm{F}) \tag{A1b}$$

$$CSI: \ \mathrm{H}/(\mathrm{H}+\mathrm{M}+\mathrm{F}) \tag{A1c}$$

$$Hit\ bias: \ (\mathrm{H}+\mathrm{F})/(\mathrm{H}+\mathrm{M}) \tag{A1d}$$

References

1. Hong, Y.; Adler, R.F.; Negri, A.; Huffman, G.J. Flood and landslide applications of near real-time satellite rainfall products. *Nat. Hazards* **2007**, *43*, 285–294. [CrossRef]
2. Hossain, F.; Anagnostou, E.N. Assessment of current passive-microwave- and infrared-based satellite rainfall remote sensing for flood prediction. *J. Geophys. Res. Atmos.* **2004**, *109*, D07102. [CrossRef]
3. Johnson, G.E.; Achutuni, V.R.; Thiruvengadachari, S.; Kogan, F. The Role of NOAA Satellite Data in Drought Early Warning and Monitoring: Selected Case Studies. In *Drought Assessment, Management, and Planning: Theory and Case Studies*; Natural Resource Management and Policy; Springer: Boston, MA, USA, 1993; pp. 31–47, ISBN 978-1-4613-6416-0.
4. Krajewski, W.F.; Anderson, M.C.; Eichinger, W.E.; Entekhabi, D.; Hornbuckle, B.K.; Houser, P.R.; Katul, G.G.; Kustas, W.P.; Norman, J.M.; Peters-Lidard, C. A remote sensing observatory for hydrologic sciences: A genesis for scaling to continental hydrology. *Water Resour. Res.* **2006**, *42*. [CrossRef]
5. 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]
6. 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]

7. Bumke, K.; Fennig, K.; Strehz, A.; Mecking, R.; Schröder, M. HOAPS precipitation validation with ship-borne rain gauge measurements over the Baltic Sea. *Tellus Dyn. Meteorol. Oceanogr.* **2012**, *64*, 18486. [CrossRef]
8. Skofronick-Jackson, G.; Kirschbaum, D.; Petersen, W.; Huffman, G.; Kidd, C.; Stocker, E.; Kakar, R. The Global Precipitation Measurement (GPM) mission's scientific achievements and societal contributions: Reviewing four years of advanced rain and snow observations. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 27–48. [CrossRef]
9. Anagnostou, E.N.; Krajewski, W.F.; Smith, J. Uncertainty quantification of mean-areal radar-rainfall estimates. *J. Atmos. Ocean. Technol.* **1999**, *16*, 206–215. [CrossRef]
10. Kent, E.; Hall, A.D.; Leader, V.T.T. The voluntary observing ship (VOS) scheme. In Proceedings of the 2010 AGU Ocean Sciences Meeting, Portland, OR, USA, 22–26 February 2010; American Geophysical Union: Washington, DC, USA.
11. Smith, S.R.; Rettig, J.; Rolph, J.; Hu, J.; Kent, E.C.; Schulz, E.; Verein, R.; Rutz, S.; Paver, C. The data management system for the Shipboard Automated Meteorological and Oceanographic System (SAMOS) initiative. In Proceedings of the OceanObs' 09: Sustained Ocean Observations and Information for Society Conference, Venice, Italy, 21–25 September 2009; Volume 2.
12. Hayes, S.P.; Mangum, L.J.; Picaut, J.; Sumi, A.; Takeuchi, K. TOGA-TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bull. Am. Meteorol. Soc.* **1991**, *72*, 339–347. [CrossRef]
13. Serra, Y.L. Precipitation measurements from the Tropical Moored Array: A review and look ahead. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 221–234. [CrossRef]
14. Bowman, K.P. Comparison of TRMM precipitation retrievals with rain gauge data from ocean buoys. *J. Clim.* **2005**, *18*, 178–190. [CrossRef]
15. Serra, Y.L.; McPhaden, M.J. In situ observations of diurnal variability in rainfall over the tropical Pacific and Atlantic Oceans. *J. Clim.* **2004**, *17*, 3496–3509. [CrossRef]
16. Serra, Y.L.; McPhaden, M.J. Multiple time-and space-scale comparisons of ATLAS buoy rain gauge measurements with TRMM satellite precipitation measurements. *J. Appl. Meteorol.* **2003**, *42*, 1045–1059. [CrossRef]
17. Levizzani, V.; Kidd, C.; Aonashi, K.; Bennartz, R.; Ferraro, R.R.; Huffman, G.J.; Roca, R.; Turk, F.J.; Wang, N.-Y. The activities of the International Precipitation Working Group. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 3–15. [CrossRef]
18. Klepp, C. The oceanic shipboard precipitation measurement network for surface validation—OceanRAIN. *Atmos. Res.* **2015**, *163*, 74–90. [CrossRef]
19. Kucera, P.; Klepp, C. Validation of High Resolution IMERG Satellite Precipitation over the Global Oceans using OceanRAIN. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 23–28 April 2017; Volume 19, p. 11794.
20. Burdanowitz, J.; Klepp, C.; Bakan, S.; Buehler, S.A. Towards an along-track validation of HOAPS precipitation using OceanRAIN optical disdrometer data over the Atlantic Ocean. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 235–254. [CrossRef]
21. 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 Theoretical Basis Document (ATBD) Version 4.5; NASA: Greenbelt, MD, USA, 2015; 30p.
22. Iguchi, T.; Kozu, T.; Meneghini, R.; Awaka, J.; Okamoto, K. Rain-profiling algorithm for the TRMM precipitation radar. *J. Appl. Meteorol.* **2000**, *39*, 2038–2052. [CrossRef]
23. Iguchi, T.; Kozu, T.; Kwiatkowski, J.; Meneghini, R.; Awaka, J.; Okamoto, K. 'ichi Uncertainties in the rain profiling algorithm for the TRMM precipitation radar. *J. Meteorol. Soc. Jpn. Ser. II* **2009**, *87*, 1–30. [CrossRef]
24. Wolff, D.B.; Fisher, B.L. Comparisons of instantaneous TRMM ground validation and satellite rain-rate estimates at different spatial scales. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 2215–2237. [CrossRef]
25. Yang, S.; Olson, W.S.; Wang, J.-J.; Bell, T.L.; Smith, E.A.; Kummerow, C.D. Precipitation and latent heating distributions from satellite passive microwave radiometry. Part II: Evaluation of estimates using independent data. *J. Appl. Meteorol. Climatol.* **2006**, *45*, 721–739. [CrossRef]
26. Petersen, W.; Kirstetter, P.; Wolff, D.; Kidd, C.; Tokay, A.; Chandrasekar, V.; Gatlin, P.; Tan, J.; Grecu, M.; Huffman, G. GPM Level 1 Science Requirements: Science and Performance Viewed from the Ground. In Proceedings of the 8th International Precipitation Working Group Workshop, Bologna, Italy, 3–7 October 2016.
27. Lempio, G.E.; Bumke, K. Measurements of solid precipitation with an optical disdrometer. *Int. BALTEX Secr. Publ. Ser.* **2007**, *38*, 123–124. [CrossRef]

28. Klepp, C.; Michel, S.; Protat, A.; Burdanowitz, J.; Albern, N.; Louf, V.; Bakan, S.; Dahl, A.; Thiele, T. *Ocean Rainfall and Ice-Phase Precipitation Measurement Network—OceanRAIN-M*; World Data Center for Climate (WDCC) at DKRZ: Hamburg, Germany, 2017.
29. Klepp, C.; Michel, S.; Protat, A.; Burdanowitz, J.; Albern, N.; Kähnert, M.; Dahl, A.; Louf, V.; Bakan, S.; Buehler, S.A. OceanRAIN, a new in-situ shipboard global ocean surface-reference dataset of all water cycle components. *Sci. Data* **2018**, *5*, 180122. [CrossRef] [PubMed]
30. Roebeling, R.A.; Wolters, E.L.A.; Meirink, J.F.; Leijnse, H. Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data. *J. Hydrometeorol.* **2012**, *13*, 1552–1566. [CrossRef]
31. McColl, K.A.; Vogelzang, J.; Konings, A.G.; Entekhabi, D.; Piles, M.; Stoffelen, A. Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophys. Res. Lett.* **2014**, *41*, 6229–6236. [CrossRef]
32. Alemohammad, S.H.; McColl, K.A.; Konings, A.G.; Entekhabi, D.; Stoffelen, A. Characterization of precipitation product errors across the United States using multiplicative triple collocation. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 3489–3503. [CrossRef]
33. Stoffelen, A. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *J. Geophys. Res. Oceans* **1998**, *103*, 7755–7766. [CrossRef]
34. Gruber, A.; Su, C.-H.; Zwieback, S.; Crow, W.; Dorigo, W.; Wagner, W. Recent advances in (soil moisture) triple collocation analysis. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 200–211. [CrossRef]
35. Chakraborty, A.; Kumar, R.; Stoffelen, A. Validation of ocean surface winds from the OCEANSAT-2 scatterometer using triple collocation. *Remote Sens. Lett.* **2013**, *4*, 84–93. [CrossRef]
36. Dorigo, W.; Scipal, K.; Parinussa, R.M.; Liu, Y.Y.; Wagner, W.; De Jeu, R.A.; Naeimi, V. Error Characterisation of Global Active and Passive Microwave Soil Moisture Data Sets. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 2605–2616. [CrossRef]
37. Loew, A.; Schlenz, F. A dynamic approach for evaluating coarse scale satellite soil moisture products. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 75–90. [CrossRef]
38. Zwieback, S.; Scipal, K.; Dorigo, W.; Wagner, W. Structural and statistical properties of the collocation technique for error characterization. *Nonlinear Process. Geophys.* **2012**, *19*, 69–80. [CrossRef]
39. 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]
40. Roebber, P.J. Visualizing multiple measures of forecast quality. *Weather Forecast.* **2009**, *24*, 601–608. [CrossRef]
41. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]
42. Schaefer, J.T. The critical success index as an indicator of warning skill. *Weather Forecast.* **1990**, *5*, 570–575. [CrossRef]
43. NASA GISS: Science Briefs: Cloud Climatology: Distribution and Character of Clouds. Available online: https://www.giss.nasa.gov/research/briefs/rossow_01/distrib.html (accessed on 16 November 2018).
44. Adler, R.F.; Negri, A.J. A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteorol.* **1988**, *27*, 30–51. [CrossRef]
45. 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]
46. 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]
47. Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]

48. Khan, S.; Maggioni, V.; Kirstetter, P.-E. Investigating the Potential of Using Satellite-Based Precipitation Radars as Reference for Evaluating Multisatellite Merged Products. *J. Geophys. Res. Atmos.* **2018**, *123*, 8646–8660. [CrossRef]
49. 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]

remote sensing

MDPI

Article

Potential of Cost-Efficient Single Frequency GNSS Receivers for Water Vapor Monitoring

Andreas Krietemeyer [1,*], Marie-claire ten Veldhuis [1], Hans van der Marel [1], Eugenio Realini [2] and Nick van de Giesen [1]

[1] Faculty of Civil Engineering, TU Delft, 2628 CN Delft, The Netherlands; J.A.E.tenVeldhuis@tudelft.nl (M.-c.t.V.); H.vanderMarel@tudelft.nl (H.v.d.M.); N.C.vandeGiesen@tudelft.nl (N.v.d.G.)

[2] GReD—Geomatics Research & Development srl, 22074 Lomazzo (CO), Italy; eugenio.realini@g-red.eu

* Correspondence: A.Krietemeyer@tudelft.nl

Received: 17 August 2018; Accepted: 13 September 2018; Published: 18 September 2018

Abstract: Dual-frequency Global Navigation Satellite Systems (GNSSs) enable the estimation of Zenith Tropospheric Delay (ZTD) which can be converted to Precipitable Water Vapor (PWV). The density of existing GNSS monitoring networks is insufficient to capture small-scale water vapor variations that are especially important for extreme weather forecasting. A densification with geodetic-grade dual-frequency receivers is not economically feasible. Cost-efficient single-frequency receivers offer a possible alternative. This paper studies the feasibility of using low-cost receivers to increase the density of GNSS networks for retrieval of PWV. We processed one year of GNSS data from an IGS station and two co-located single-frequency stations. Additionally, in another experiment, the Radio Frequency (RF) signal from a geodetic-grade dual-frequency antenna was split to a geodetic receiver and two low-cost receivers. To process the single-frequency observations in Precise Point Positioning (PPP) mode, we apply the Satellite-specific Epoch-differenced Ionospheric Delay (SEID) model using two different reference network configurations of 50–80 km and 200–300 km mean station distances, respectively. Our research setup can distinguish between the antenna, ionospheric interpolation, and software-related impacts on the quality of PWV retrievals. The study shows that single-frequency GNSS receivers can achieve a quality similar to that of geodetic receivers in terms of RMSE for ZTD estimations. We demonstrate that modeling of the ionosphere and the antenna type are the main sources influencing the ZTD precision.

Keywords: GNSS meteorology; GPS; Zenith Tropospheric Delay; precipitable water vapor; SEID; single frequency GNSS; Precise Point Positioning; low-cost receivers; goGPS

1. Introduction

Water vapor plays an important role for atmospheric processes. It is the most abundant greenhouse gas [1] and is spatially and temporally highly variable [2]. Atmospheric water vapor is essential for convection in the lower atmosphere and hence crucial for the generation of clouds and rainfall. A relationship between water vapor fields and severe weather events has been observed, e.g., by Seko et al. [3]. With regard to a warming climate, the amount of water vapor in the atmosphere increases and causes additional absorption of long-wave radiation and reflects it back to the ground [4]. Water vapor also transports latent heat through the atmosphere. Distribution of atmospheric water vapor is an important factor for weather models, and its monitoring is crucial for weather research. The traditional method to measure water vapor in the air is by releasing radiosonde balloons. However, the releases are typically only performed a few times per week and are characterized by distances often greater than 100 km.

Another approach is to determine Precipitable Water Vapor (PWV) from the observed delay in Global Navigation Satellite System (GNSS) transmitted signals. The estimated Zenith Tropospheric

Delay (ZTD) is the driving factor to obtain PWV, and uncertainties in ZTD dominate the error budget of PWV [5]. Monitoring local inhomogeneities of water vapor that may be associated with deep convection is important for the prediction of severe weather events. One should note that existing weather radars can only detect clouds after the start of precipitation. Currently, PWV derived from globally, nationally, and regionally distributed networks of dual-frequency receivers is used to improve existing numerical weather models [6]. The networks were originally designed for geodetic applications and are characterized by an inhomogeneous distribution and inter-station distances in the order of tens of kilometers or above. Due to their high inter-station distances, these networks are often not sufficiently dense to monitor small-scale water vapor fluctuations ([7,8]). Real-Time Kinematic (RTK) networks are already characterized by station distances below 30 km (e.g., Netherlands or Germany). Some networks, such as the Japanese GEONET, even reach inter-station distances of about 20 km [9]. In our study, networks with average station distances of around 30 km are considered as dense. Regional networks in Europe with station distances between 70 and 100 km are considered as low-density networks. However, even networks with inter-station distances of 20 km are not dense enough to properly detect water vapor variations associated with localized convective phenomena [10]. The cost of dual-frequency receivers exceeds several thousands dollars. Driven by these economic reasons, the use of dual-frequency receivers for water vapor monitoring is not feasible at high density. Instead, inexpensive single-frequency receivers already available at a cost of a hundred to a few hundred dollars offer an interesting alternative. Having been designed to be embedded into mass-market devices such as car navigation equipment and smartphones, these inexpensive receiver modules and antennas are more economical than their dual-frequency geodetic counterparts in terms of both money and power consumption (in case of off-grid application).

Currently, widespread application of single-frequency GNSS is limited by errors introduced by the ionosphere. Different strategies exist to obtain ZTD estimations, e.g., the network approach based on double differences or the Precise Point Positioning (PPP) based on undifferenced observations. A comprehensive analysis using different mapping functions, cutoff angles, and positioning methods has been performed by Baldysz et al. [11]. In that study, the PPP performed better than the differential approach for PWV linear trend analysis. To account for the error caused by the ionosphere, we apply the Satellite-specific and Epoch-differenced Ionospheric Delay (SEID) model as proposed by Deng et al. [12]. This allows us to treat the single-frequency data as dual-frequency data and process it with well known PPP software [13]. The primary goal of this study is to investigate the feasibility of low-cost single-frequency receivers to densify existing dual-frequency networks and examine the error sources that the receivers are prone to. The research is a first step to examine the potential wide-spread deployment of low-cost GNSS receivers in data sparse regions with limited investment budgets. It aims to highlight the trade-off between using cost-efficient devices compared to their expensive, geodetic-grade dual-frequency counterparts for water vapor monitoring. For this purpose, we conducted a monitoring experiment alongside an International GNSS Service (IGS) station and compared our results with available reference datasets. Earlier studies, e.g., by Deng et al. [12], Deng et al. [14], and Realini et al. [8] demonstrated the feasibility of applying single-frequency GNSS in PPP mode. In our study, we used two different types of low-cost on-site single-frequency receivers and investigated their performance for monitoring water vapor. Durable, weather-proof stations incorporating low-cost GNSS receivers are a recent development and were not available on the market before 2015. Even today only a few startup and small- and medium-sized enterprises produce them. One of the low-cost single-frequency stations used in this study is the GeoGuard Monitoring Unit (GMU) designed by the Italian company GReD [15]. All components are Ingress Protection (IP) 66 or 67 certified and the first units operate since November 2015. Currently 70 GMU stations are continuously monitored by GeoGuard from which some are also deployed in harsh conditions, such as mountainous areas above 2000 m and near the sea. Even though our experimental setup could make use of an already dense RTK network for the ionospheric interpolation, we selected more challenging station distances and only a minimum of SEID reference stations. The study uses

exclusively open-source models and publicly available datasets to support the applicability of the setup to other locations. By co-locating multiple different sensors and splitting the antenna signal of a calibrated geodetic-grade antenna to a geodetic-grade receiver, low-cost single- and dual-frequency devices, we were able to make statements about the impact of receiver-and antenna-related errors. The study is organized as follows: Section 2 describes the basics of the GNSS meteorology, the SEID algorithm, and the experimental setup and data processing methods. Section 3 illustrates the ZTD reference comparison, the experimental setup results using two different SEID constellations, and the split antenna experiment results. Sections 4 and 5 contain the discussion and conclusions, respectively.

2. Methodology

2.1. Water Vapor from GNSS Measurements

Traditional GNSS geodesy aims to obtain precise positioning information from the signal delay between a GNSS receiver and satellites in sight. The delay is affected by several error sources such as clock errors, antenna phase center variations, and tropospheric and ionospheric delays. The tropospheric delay is further separated into a dry (hydrostatic) delay and a wet delay as a result of atmospheric water vapor. The Slant Total Delays (STDs) along each satellite-receiver line of sight are mapped to the zenith direction in order to estimate the ZTD. In GNSS meteorology, the ZTD can be directly assimilated in numerical weather models, or the ZTD can be converted into PWV using surface pressure data and a simple model of the atmosphere.

According to Thayer [16], the refractivity in the troposphere can be considered as the sum of a dry (hydrostatic) and wet component, which can be related to the atmospheric temperature and partial pressure of water vapor and dry gases:

$$N = N_{dry} + N_{wet} = k_1 \frac{p_d}{T} + k_2 \frac{e}{T} + k_3 \frac{e}{T^2} \tag{1}$$

where $k_1 \frac{p_d}{T}$ accounts for the dry part and the last two terms for the wet part. p_d is the combined partial pressure of dry gases in mbar, T the temperature in degree Kelvin, and e the partial pressure of water vapor in mbar. The empirical constants $k_1 = (77.604 \pm 0.014) K \ \text{mbar}^{-1}$, $k_2 = (64.79 \pm 0.08) K$ mbar^{-1} and $k_3 = (3.776 \pm 0.004) 10^5 K^2 \ \text{mbar}^{-1}$ were determined by Thayer [16]. The integrals of the refractivity N in the zenith direction are referred to as Zenith Hydrostatic Delay (ZHD) and Zenith Wet Delay (ZWD). The ZHD can be estimated using the Saastamoinen model [17]:

$$\text{ZHD} = \frac{0.0022768 P_0}{1 - 0.00266 \cos(2\Phi) - 0.00028 H} \tag{2}$$

where P_0 is the total atmospheric pressure in mbar expected to be observed at the receiver position, Φ is the latitude in radians and H the orthometric height (i.e., above the reference geoid) in kilometers.

Since the wet delay is much more variable than the hydrostatic delay and its predictive value is poor in comparison to the ZHD estimation, the difference between the ZTD and the modeled ZHD is used to obtain the ZWD. The authors of Bevis et al. [18] demonstrate how the ZTD observed from GPS measurements can be used to derive PWV with given pressure and temperature data. Using the ZWD, the PWV can be estimated with a non-dimensional ratio Π:

$$\Pi = \frac{10^6}{\rho R_v [(\frac{k_3}{T_m}) + k_2']} \tag{3}$$

with T_m being the weighted mean temperature of the atmosphere in degree Kelvin, ρ the density of liquid water in kg m^{-3}, R_v the specific gas constant of water vapor in J $(\text{kg K})^{-1}$, and k_2' defined as $k_2' = k_2 - m k_1$ with m being the ratio of molar masses of water vapor (18.015 g mol^{-1}) and dry air (28.964 g mol^{-1}). The constants k_1, k_2, and k_3 are the same as proposed byThayer [16]. The weighted mean temperature of the atmosphere T_m is computed based on radiosonde measurements. However,

a simplified model using surface temperature measurements is sufficiently accurate. In our analysis, we use the T_m derived by Baltink et al. [19] based on 9129 radiosonde ascents in De Bilt, The Netherlands:

$$T_m = 0.673 T_s + 83.0 \tag{4}$$

with T_s defined as the surface temperature in Kelvin. The water vapor conversion factor Π ranges typically around 0.16 and varies up to 20% depending on the location, height, and meteorological conditions [20]. The PWV is related to the Integrated Water Vapor (IWV) overlying the receiver, whereby the PWV is defined as the height of an equivalent column of liquid water. The PWV is obtained by

$$\mathrm{PWV} = \Pi \cdot \mathrm{ZWD}. \tag{5}$$

Since there are no on-site temperature and air pressure observations available at our experimental setup, we are using measurements from the nearest weather station at a 5 km distance and with a negligible elevation difference. We utilize the EGM2008 geoid model [21] to convert between the ellipsoidal and orthometric height at the station position. The conversion between mean-sea-level atmospheric pressure to the station height is performed by applying a standard lapse rate formula in mbar per meter after the ideal gas law [22]:

$$\frac{\delta P}{\delta h} = -\frac{P\, M\, g}{R\, (T + 273.2)} \tag{6}$$

with P being the pressure in mbar, T the temperature in degree Celsius, M the mean molar mass of atmospheric gases (0.02896 kg mol^{-1}), g the gravity acceleration (9.807 m s^{-2}), and R the universal gas constant (8.314 J K^{-1} mol^{-1}), respectively. With P being the pressure in mbar and T the temperature in degree Celsius. Accounting for the geoidal undulation is important since the height error will develop a pressure offset. Utilizing the Saastamoinen formula (Equation (2)) shows that a pressure error of one mbar results in an error of around 2.28 mm in ZHD at our station coordinates.

2.2. SEID Ionospheric Delay Modeling

Since it is not possible to obtain an ionosphere-free linear combination of sufficient accuracy based on single-frequency receiver data, the ionosphere must be modeled. In order to process single-frequency receivers within a network of dual-frequency receivers with conventional GNSS processing engines, the authors of Deng et al. [12] developed the SEID model. They use the geometry-free ionospheric observation L_4 [23], which is defined as

$$L_4 = L_1 - L_2 = \lambda_1 A_1 - \lambda_2 A_2 - 40.28\left(\frac{1}{f_1^2} - \frac{1}{f_2^2}\right) \cdot STEC(\lambda, \theta). \tag{7}$$

For simplicity the noise and multipath errors are ignored in the equation. L_1 and L_2 denote the carrier-phase measurements. The term $40.28(\frac{1}{f_1^2} - \frac{1}{f_2^2}) \cdot STEC(\lambda, \theta)$ denotes the ionospheric delay on the system-dependent frequencies f_1 and f_2. $STEC$ is the Slant Total Electron Content, which is defined as the integral of the electron density along the signal path. λ and θ are latitude and longitude of the Ionospheric Pierce Points (IPPs) of the ray paths on the single-shell defined ionosphere layer at a 350 km height. λ_1 and λ_2 are the wavelengths and A_1 and A_2 denote the phase ambiguities on the respective frequencies. Ref. [24] demonstrate that the epoch-differenced ionospheric delays are sufficiently accurate to remove the ionospheric delay under normal ionospheric conditions for ZTD estimations. Since the estimation of the ambiguity parameter remains a major obstacle in GNSS processing, the epoch-differences dL_4 are used:

$$dL_4(j+1) = L_4(j+1) - L_4(j) = 40.28(\frac{1}{f_1^2} - \frac{1}{f_2^2}) \cdot dSTEC(\lambda, \theta). \quad (8)$$

Utilizing this equation, the ambiguity parameters between two continuously tracked epochs is eliminated. One should note that cycle slips must be removed and phase center variations applied in advance. In small-scale networks, the $dSTEC$ can be modeled with a linear function with the latitude and longitude θ and λ of the IPPs. Each pair of epochs and each satellite can be approximated by the linear function

$$dL_4(j+1) = a_0 + a_1\lambda + a_2\theta \quad (9)$$

where the model parameters a_0, a_1, and a_2 are to be estimated by means of least-squares from the epoch-differenced ionospheric delays of the surrounding dual-frequency stations. Using this model, the epoch-differenced ionospheric correction $d\tilde{L}_4(j+1)$ of any receiver within the spanned network may be calculated. The sum of the continuously tracked ionospheric corrections from j_0 to j_n results in

$$\tilde{L}_4(j_n) = \sum_{j_0}^{j_n} d\tilde{L}_4 + L_4(j_0). \quad (10)$$

As depicted in Equation (10), the ionospheric delay at epoch j_0 is unknown (thus set to zero). Unfortunately, this destroys the integer nature of the phase ambiguity in the interpolated data, meaning it can never be fixed to integer. However, since in typical PPP, the ambiguities are not fixed to integer (but estimated as float values), this has no impact on the data processing. The synthesized $\tilde{L}_2$ signal is constructed from the difference between L_1 and $\tilde{L}_4$:

$$\tilde{L}_2 = L_1 - \tilde{L}_4. \quad (11)$$

The synthesized observable provides the same information as L_1, except that the ionospheric delay is adjusted to the L_2 frequency. Employing the generated dual-frequency data allows the data to be processed in PPP-mode in conventional PPP analysis tools.

2.3. Experimental Setup and Data Processing

The GNSS data server at the TU Delft collects data from the Dutch Permanent GNSS Array (DPGA), which consists of up to 18 continuously operating dual-frequency receivers throughout the Netherlands. The regional network density is considered as low since its stations are typically located at a 50–100 km distance. The data is publicly available and mainly used for research and educational purposes and providing data to the International GNSS Service (IGS), EUREF Permanent Network (EPN), and EUMETNET GNSS water vapour programme (E-GVAP).

To investigate the feasibility of low-cost single-frequency receivers for water vapor monitoring, we designed an experimental setup along the existing IGS station DLF1 in Delft, The Netherlands. The setup consists of one dual-frequency receiver (IGS station DLF1) and two co-located low-cost single-frequency receivers. The IGS station DLF1 is equipped with a geodetic-grade calibrated choke ring antenna. The antenna type is the current state-of-the-art and serves in our experiment as a reference for the best available antenna observations. One single-frequency station is the GMU which consists of a u-blox LEA-M8T single-frequency receiver with a Tallysman TW3470 antenna. The device has been successfully validated and is currently used for operational deformation monitoring [15]. It is fully autonomous and equipped with a solar panel and wireless (3G) data communication. Additionally, we co-located a u-blox NEO-M8T evaluation toolkit (UBX) equipped with its standard ANN-MS patch antenna.

To apply the SEID algorithm, we selected two SEID reference constellations. The first constellation (DPGA case) consists of 4 stations (VLIS, IJMU, KOS1, and APEL) from the Dutch permanent network.

The stations KOS1 and APEL were chosen as a backup for each other, since both have data gaps during the observation period. These are, however, different dates and the SEID model is still functioning by using only 3 instead of 4 reference stations. The total area covered in this case is about 6400 km^2. To analyze the effect of the SEID reference network distances, we added a second SEID constellation (EUREF case) utilizing stations from the EUREF network to our study. The stations are HERT (Great Britain), BORJ (Germany), and DOUR (Belgium). They are located approximately 305, 240, and 215 km from the experimental setup in Delft. The conversion from ZTD to PWV was performed by using data from an Automatic Weather Station (AWS) located 5 km from Delft, operated by the Royal Netherlands Meteorological Institute (KNMI) for temperature and air pressure observations. The station locations are illustrated in Figure 1.

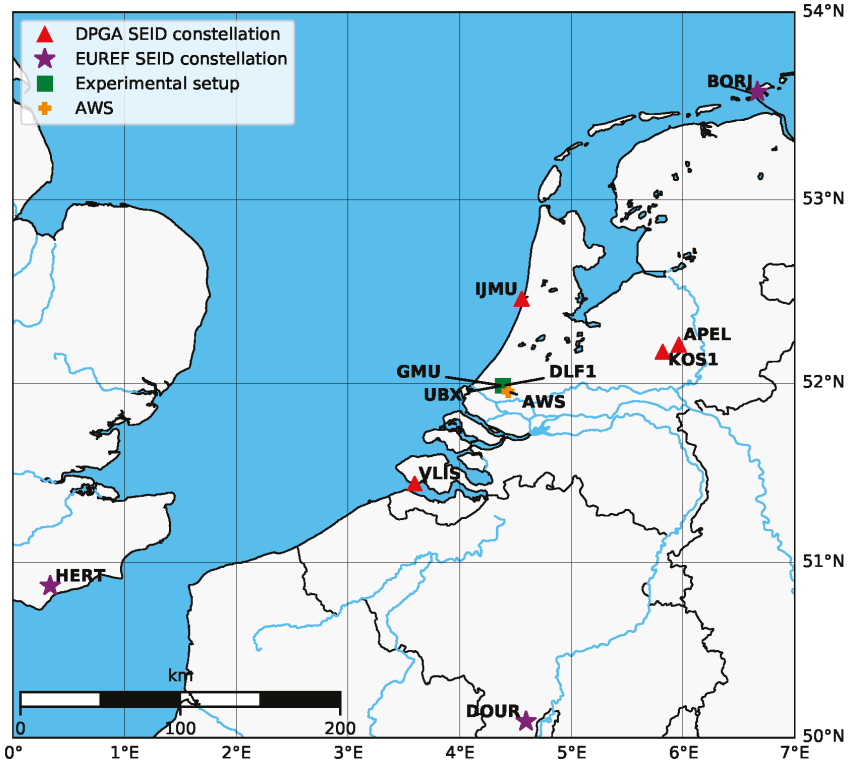


Figure 1. Positions of the SEID reference stations of the DPGA (red triangles) and EUREF (purple stars) constellations, the experimental setup in Delft (green square), and the AWS (orange cross). The sites DLF1, Geo Monitoring Unit (GMU), and u-blox NEO-M8T evaluation toolkit (UBX) are all co-located within a 10 m^2 area.

We selected the analysis period from 1 January until 31 December 2017 to cover both low (winter) and high (summer) water vapor activities. During winter, the PWV at this location typically ranges between 1 and 30 mm, whereas the summer is characterized with PWV values between 10 and 50 mm. However, the single-frequency units GMU and UBX have only provided data since September 2017. The IGS station DLF1 provides continuous high-rate data throughout the year with a temporal resolution of up to 1 s. For our analysis, we use this dual-frequency data to perform PPP to obtain a reference ZTD dataset. In order to validate the SEID modeling, we simulate a single-frequency unit

(DLF1_SF) by removing the code and phase data from the second frequency (C2 and L2) with the teqc [25] software. Consequently, DLF1 in dual-frequency (DLF1_DF) and synthesized single-frequency mode (DLF1_SF) provide data throughout the year. For all sites (including the SEID network stations), we use 30-second data in the Receiver Independent Exchange Format (RINEX).

The SEID processing and subsequent PPP processing are performed by using the open-source MATLAB program goGPS version 0.4.3 [26]. This version is based on a forward Kalman filter approach. During the pre-processing step, cycle slips are corrected and antenna calibrations, if available, are applied to all stations. An epoch-by-epoch code-only least-squares adjustment is then executed to identify and remove outliers. Observation sets (i.e., code and phase observations for a given satellite) for which the configured residual code observation error threshold is exceeded, are flagged as outliers and removed. Entire epochs are removed if the number of available observation sets is equal or lower than 4, or if the standard deviation of the estimated error exceeds a given threshold. During the PPP processing, an additional outlier detection and removal algorithm is executed, based on a configured threshold on phase post-fit residuals. In this version of goGPS, the PPP Kalman filter is not reset at the day-boundary. This enables a seamless ZTD estimation without spikes at the beginning of each dataset. Both SEID constellation studies and the PPP estimations share the same configuration parameters. Table 1 summarizes the settings of our analysis.

Table 1. SEID and ZTD Computation Settings.

Item	Processing Strategies
Software	goGPS v. 0.4.3
Observations	GPS-only
Sampling interval	30-second
Processing mode	SEID-PPP
Antenna calibration	IGS (if available)
Troposphere modeling	Saastamoinen (with GPT model)
Troposphere mapping function	GMF
Elevation cutoff	10°
Ocean loading	FES2004
Observation weighting	same weight for all observations
Clock & orbits	IGS Final
Kalman filter reset	no (seamless)
Code observation error threshold	30 m
Phase observation error threshold	0.05 m
Code least-squares estimation error st. dev. threshold	40 m

After the SEID interpolation on the single-frequency datasets, we obtain dual-frequency RINEX files which we process in PPP mode in goGPS. We use data from GPS-only satellites, an elevation cutoff angle of 10 degrees and apply ocean loading effects obtained from the FES2004 model [27]. The IGS antenna calibration, final orbits, 30-second satellite clocks, and earth rotation products are used [28]. One should note that the IGS final products are only available with a 12–18 day latency. Hence, they can only be used for post-processing purposes.

The ZTD estimations are affected by the antenna type and the SEID model. The best possible references for these are calibrated choke ring antennas and dual-frequency PPP estimations. To verify the performance of the goGPS PPP engine, the ZTD estimations are compared to existing reference datasets, respectively. The experimental setup provides these references and allows us to validate these aspects.

A first step in our analysis is the comparison of various available ZTD reference datasets from different analysis centers to results obtained for our experimental setup. Specifically, we compare the reference ZTD data from the IGS available from the Crustal Dynamics Data Information System (CDDIS) archive (IGS [29]), the Nevada Geodetic Laboratory (NGL [30]), and the EPN analysis

center outputs [31] from the Federal Agency for Cartography and Geodesy (BKG), Federal Office of Topography (LPT), and the Royal Observatory of Belgium (ROB).

3. Results

3.1. Inter-Comparison of Different ZTD Reference Datasets

Analysis Centers (ACs) provide ZTD estimations for IGS stations worldwide. However, only the IGS and NGL datasets have a 5 min temporal resolution. The EPN products ROB, LPT, and BKG are available in 1 h intervals. The reference datasets IGS, BKG, LPT, and ROB are based on the Bernese processing engine using the least-squares adjustment method. NGL provides GIPSY/OASIS II ZTD estimations in PPP using a modified Kalman filter. The IGS data is available from the CDDIS archive. The IGS AC updated their processing engine on DOY 29 in 2017 to Bernese 5.2. Table 2 summarizes the characteristics of each reference dataset.

Table 2. Characteristics of reference datasets of the IGS station DLF1. All products are post-processed results using the final satellite orbit and clock data.

AC	Processing Engine	Processing Method	Cutoff (°)	Resolution	Missing Days
IGS	Bernese 5.0 & 5.2	PPP	7	5 min	10
NGL	GIPSY/OASIS II	PPP	7	5 min	10
BKG	Bernese 5.2	Double-Differences	3	60 min	1
LPT	Bernese 5.3	Double-Differences	3	60 min	4
ROB	Bernese 5.2	Double-Differences	3	60 min	4

Surprisingly not all analysis centers provide data for all dates, whereas the IGS station DLF1 recorded data throughout the selected observation period. As a result, our PPP-processed dual- and simulated single-frequency data of DLF1 is characterized by no data gaps. The IGS and NGL datasets show 10 missing days, whilst BKG, LPT, and ROB show only 1 and 4, respectively. It must be noted that these are for the most part not the same dates. We assess the performance of the ZTD estimations by comparing the various reference datasets to each other. The scatter plots of all reference datasets, their corresponding root mean square errors (RMSE), mean biases and standard deviations are depicted in Figure 2.

The scatter plots show an RMSE of about 3 mm between the post-processed datasets IGS, NGL, BKG, LPT, and ROB. No remarkable bias is observed. The best agreement is found between the 1 h datasets BKG, ROB, and LPT. This is explained by the fact that for the 1 h datasets individual estimates are based on more measurements compared to the 5 min datasets and by the fact that the three ACs use a similar processing method. For the comparison with IGS and NGL, we use the timestamps of the 1 h datasets. The precision of the ZTD estimations between the analysis center outputs are very similar, in the order of 2–4 mm. For this reason, we only use the IGS data as a reference to compare the results from our experimental setup. The advantage of the IGS estimations is the 5 min sampling interval and the high data availability.

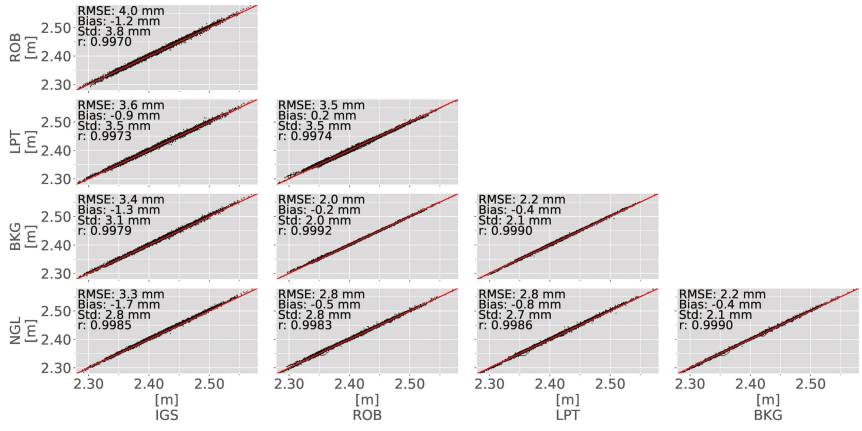


Figure 2. ZTD scatter plots of all post-processed reference datasets at DLF1 in 2017. The solid red line depicts the regression line. RMSE, bias, and std values are in millimeter. The r-value depicts the correlation.

3.2. SEID-PPP-Processed ZTD Estimations

After validating the reference data, we analyzed the results of the goGPS PPP estimations at the experimental setup location for both SEID constellations. Figure 3 shows the ZTD scatter plots between the DLF1 dual- (DLF1_DF) and single-frequency (DLF1_SF), GMU, UBX, and the selected IGS reference dataset using the DPGA SEID constellation.

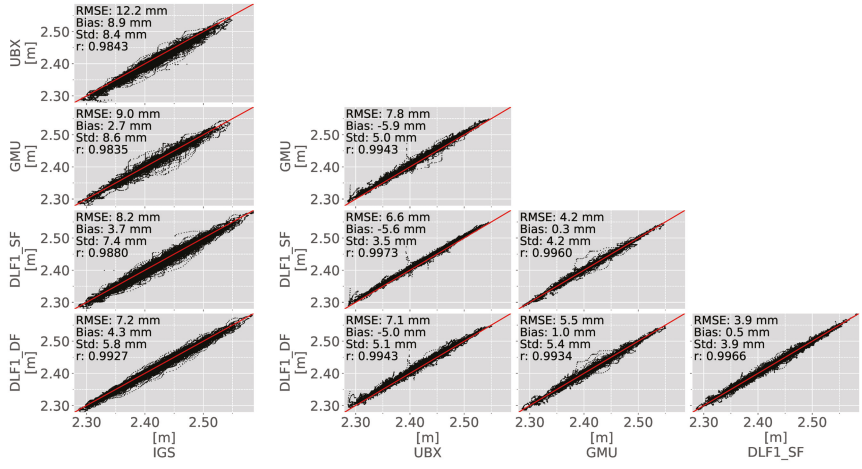


Figure 3. ZTD scatter plots of the IGS reference dataset, the dual-frequency PPP-processed DLF1 dataset (DLF1_DF), and the SEID-processed single-frequency data from UBX, GMU, and DLF1_SF utilizing the DPGA SEID reference network. The plots on the left depict the comparison of the goGPS estimations to the IGS reference dataset. The scatter plots on the right show the SEID and instrument errors. The observation period is from 1 January 2017 (DLF1_DF & DLF1_SF) and 1 September 2017 (GMU & UBX) until 31 December 2017. The solid red line depicts the regression line.

The column on the left in Figure 3 demonstrates the goGPS PPP software error by comparing the estimations to the IGS reference ZTD dataset. Compared to the IGS reference dataset, the goGPS PPP

dual-frequency dataset (DLF1_DF) shows an RMSE of 7 mm and a bias of 4 mm over the whole year. A bias in the range of 3–9 mm is present in the goGPS-processed results and negatively influences the RMSE. The scatter plots on the right show the comparisons of the goGPS estimated results among each other. The dual-frequency data (DLF1_DF) and the simulated SEID-processed single-frequency data (DLF1_SF) show an RMSE of 4 mm. The GMU yields an RMSE of 6 mm and no notable bias. The UBX dataset is largely characterized by a negative 5 mm bias but a similar (5 mm) standard deviation and slightly higher RMSE (7 mm). All goGPS-processed datasets yield standard deviations between 4 and 5 mm using the DPGA SEID constellation. Figure 4 presents the scatter plots utilizing the EUREF SEID constellation with stations characterized by distances of about 200–300 km.

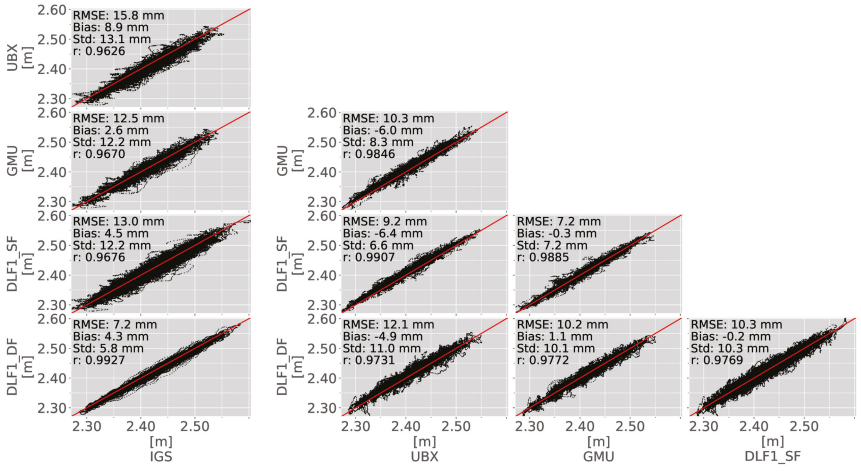
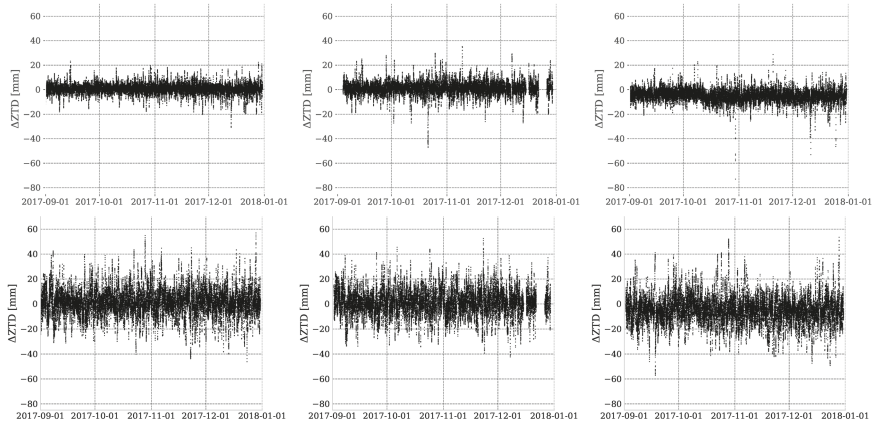


Figure 4. ZTD scatter plots of the reference dataset IGS, the dual-frequency PPP-processed DLF1 dataset, and the SEID-processed single-frequency data from UBX, GMU, and DLF1_SF utilizing the EUREF SEID constellation. The plots on the left depict the comparison of the goGPS dual-frequency estimation to the IGS reference dataset. The scatter plots on the right show the SEID and instrument errors. The observation period is from 1 September 2017 until 31 December 2017. The solid red line depicts the regression line.

Similar to the layout in Figure 3, Figure 4 depicts the software error on the left and the experiment results on the right. The bias introduced by the goGPS software remained the same. The RMSE and standard deviation increased to 13/12, 12/12, and 16/13 mm, respectively, for DLF1_SF, GMU and UBX. A higher scattering effect of the single-frequency datasets (DLF1_SF, GMU, and UBX) is evident after the SEID interpolation. Compared to the DLF1_DF reference, both GMU and DLF1_SF are now characterized by an RMSE of 10 mm. UBX shows an RMSE of 12 mm. The bias of the dataset has not changed and is still present. Among each other, the single-frequency datasets yield standard deviations between 7 and 8 mm. Compared to the dual-frequency reference, the standard deviation increased to 10–11 mm. Figure 5a–c shows the ZTD differences of the SEID-processed DLF1_SF, GMU, and UBX data compared to the reference dual-frequency dataset for the period 1 September–31 December 2017.

The differences using the DPGA SEID constellations (top plots in Figure 5) generally range between 0 and 10 mm in the DLF1_SF and GMU case. More outliers are visible and generally increased by about 10 mm for the GMU (b) compared to the DLF1_SF (a) case. As depicted in the scatter plots, a clear negative bias is seen for the UBX differences. Contrary to the synthesized dataset (DLF1_SF), the GMU and UBX datasets show an increased amplitude and an amount of outliers. The results utilizing the larger SEID reference station distances (bottom plots in Figure 5) demonstrate an increased dispersion throughout the comparison period for all sites. Except slightly higher outliers and the

previously mentioned bias in the UBX data, no performance decrease is visually evident. Figure 6 shows the boxplots to indicate the variability of the single-frequency dataset differences compared to the DLF1_DF reference utilizing both SEID constellations.



(**a**) DLF1_SF (**b**) GMU (**c**) UBX

Figure 5. ZTD differences between the goGPS-processed dual-frequency reference dataset and the SEID-processed single-frequency estimations DLF1_SF, GMU, and UBX. The figures at the top depict the DPGA SEID constellation results, whilst the ones on the bottom are from the EUREF SEID constellation, respectively. DLF1_SF (**a**) is the synthesized single-frequency dataset obtained from the IGS stations DLF1. GMU (**b**) and UBX (**c**) are the two cost-efficient single-frequency receivers utilizing different antenna types but sharing the same receiver type. A comparison period of 1 September–31 December 2017 was selected since all receivers provide observations during this period.

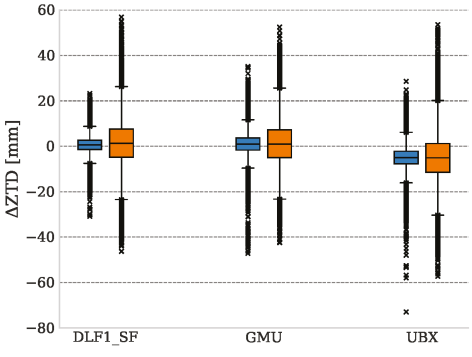
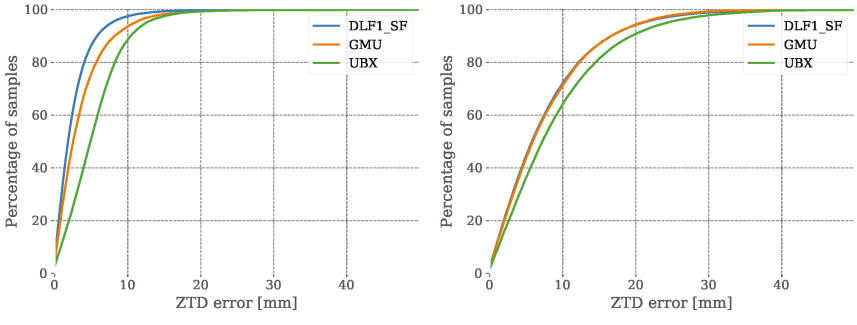


Figure 6. Boxplots of the SEID-single-frequency ZTD differences to the DLF1_DF reference using the DPGA SEID constellation (blue) and the EUREF reference stations (orange).

The figure demonstrates that the scattering effect is about the same for all stations using the same SEID constellation. Regarding only the DPGA SEID case (blue boxes), DLF1_SF shows the smallest errors. The scattering effect is also marginally decreased compared to the GMU and UBX case. However, higher outliers in the GMU and UBX datasets are evident. Consistent to the negative

UBX bias, the amount of outliers are predominately negative. Regarding the EUREF SEID case (orange boxes), the variation is rather equivalent for all stations. The bias and slightly increased outliers remain for the UBX case. However, comparing the two SEID cases among each other, a higher variation and more outliers are present. Figure 7 shows the ZTD cumulative error distributions of the single-frequency datasets (DLF1_SF, GMU, and UBX) to the dual-frequency reference dataset for both SEID cases.



(**a**) DPGA SEID case

(**b**) EUREF SEID case

Figure 7. Cumulative error distribution plots of the single-frequency ZTD estimations DLF1_SF, GMU, and UBX compared to the DLF1_DF reference dataset. The figure at the left (**a**) depicts the DPGA SEID constellation results, while (**b**) is from the EUREF SEID case, respectively.

The cumulative error distribution of the DPGA SEID case datasets is depicted in Figure 7a. The figure shows that about 95% (DLF1_SF) and about 90% (GMU) of the data points are within 10 mm precision. About 85% of the UBX data shows an error below 10 mm. Using the larger EUREF SEID station network (Figure 7b), a clear degradation is evident. Only 70% (DLF1_SF, GMU) and 60% (UBX) of the data show an error below 10 mm compared to the dual-frequency reference. About 95% of the ZTD estimations from DLF1_SF and GMU are below a 20 mm error range, which applies to about 90% of the UBX data.

To match a PWV quality requirement of 1–3 mm precision for application in NWP models [32], we selected a threshold of 10 mm ZTD (results in 1–2 mm PWV) difference as a valuable input for meteorological applications. Table 3 summarizes the statistics of the sites compared to the dual-frequency reference dataset for both SEID cases. Apart from the RMSE, the mean bias, correlation, and percentage of data points exceeding the 3σ limit and greater than 10 mm ZTD differences are listed. Another quality measure is the amount of data points that were not available, either because they were removed during data processing as outliers or because they were not recorded by the measuring station.

Table 3. Statistics on the experimental setup data using dual-frequency ZTD estimations as reference. Units of the RMSE, bias, and standard deviation (σ) are depicted in mm.

Case	Site	RMSE [mm]	Bias [mm]	σ [mm]	Corr	%$\geq 3\sigma$	%$\geq \Delta$10 mm	%Missing
DPGA	DLF1_SF	3.93	0.52	3.90	0.9967	1.64	2.61	0.46
	GMU	5.55	0.99	5.46	0.9938	1.44	6.77	7.10
	UBX	7.10	−4.96	5.08	0.9945	2.63	12.68	0.87
EUREF	DLF1_SF	10.32	−0.20	10.32	0.9769	0.92	29.04	0.28
	GMU	10.20	1.11	10.14	0.9772	0.62	29.87	6.89
	UBX	12.09	−4.93	11.04	0.9731	1.38	37.12	0.86

For the DPGA case, the SEID-processed DLF1_SF data shows 2.61% of their data points that exceed the 10 mm threshold. The GMU yields 6.77%, and UBX yields 12.68%. The simulated dataset DLF1_SF shows the best performance compared to its dual-frequency counterpart. It shows the lowest RMSE (3.93 mm), bias (0.52 mm), and standard deviation (3.90 mm) and yields the highest correlation (0.9967). The GMU is characterized by a smaller RMSE (5.55 mm) compared to the UBX station (7.10 mm). However, the correlation and standard deviation of UBX slightly outperforms the ones from the GMU. The differences are marginal and may be a coincidence. Different antenna types were used, and the GMU is also characterized by more data gaps that are probably the driving factor behind this effect. Another quality decline is observed for the percentage of points exceeding the 3σ limit comparing the GMU and the UBX data. Even though the σ values are about the same, about twice the amount exceeds the 3σ limit at the latter one. A similar effect can be seen by the amount of points that exceed the 10 mm threshold (12.68%). However, this effect is mainly caused by the identified UBX bias. For the UBX dataset, 0.87% of all data has been rejected as outliers. The GMU site demonstrated a 7.1% loss of data. Except the bias in the UBX dataset, the UBX site shows a similar standard deviation as the GMU (5.08 and 5.46 mm, respectively).

Regarding the EUREF SEID case, the 10 mm threshold is exceeded for about 29–30% of the data points at the DLF1_SF and GMU datasets. The UBX dataset is not only characterized by the highest RMSE (12.09 mm) and bias (−4.93 mm) but consists also of more than 37% of data that are above the 10 mm threshold. The amount of discarded epochs (NaN values) are almost in the same range as in the first SEID case study. Expectedly, the σ values are higher, in the range of 10–11 mm, for all single-frequency solutions. However, the amount of data above the 3σ limit decreased.

3.3. PWV Computation

Utilizing Equation (3), a surface temperature error of 10 Kelvin will result in an error of approximately 0.34 percent from the original ZWD (results into an error of up to 1 mm in PWV). Since sea level air pressure is rather homogeneous over a distance of 5 km, we also use the recorded observation at the AWS as input for the ZHD computation. The extrapolation from sea level pressure to the station height is performed using Equation (6). After subtracting the computed hydrostatic delay from the estimated ZTD, the remaining ZWD is multiplied by the conversion factor. Since the conversion to PWV is performed with the same temperature and air pressure data on the reference ZTD as on the goGPS ZTD estimations, it results into a linear relationship and does not contain any additional information than the original ZTD. Hence, the PWV plots are not shown here. Instead, the PWV RMSE, biases, and standard deviations utilizing observations from the nearby AWS for both SEID cases are summarized in Table 4.

Table 4. RMSE, bias, and standard deviation (σ) on the PWV estimation from the SEID-processed solutions using the two SEID constellations DPGA and EUREF utilizing temperature and air pressure measurements from the AWS at the Rotterdam airport (5 km distance) compared to the dual-frequency reference.

Case	Site	RMSE [mm]	Bias [mm]	σ [mm]
SEID (DPGA)	DLF1_SF	0.60	0.08	0.59
	GMU	0.85	0.17	0.84
	UBX	1.05	-0.72	0.77
SEID (EUREF)	DLF1_SF	1.61	-0.03	1.61
	GMU	1.60	0.20	1.59
	UBX	1.86	-0.70	1.72

The PWV estimations from the DPGA (smaller SEID network) show comparable results. The single-frequency stations yield RMSE values of 0.6 mm (DLF1_SF), 0.85 mm (GMU), and 1.05 mm (UBX). The PWV estimations from the EUREF SEID constellation stay below a 2 mm RMSE. The priorly identified UBX bias results in about a −0.7 mm PWV bias. To distinguish between the receiver, antenna,

and SEID-introduced error, we set up an additional experiment in Italy where we split the signal from a geodetic-grade antenna into three different receiver types.

3.4. Splitting of A Geodetic Antenna to Different Receiver Types (Italy)

From GNSS basics, we know that the receiver itself and the antenna quality introduces an error to the estimations. Calibrated antennas to correct for the phase center error improve the obtained results. To further elaborate the influence of the receiver and distinguish between the SEID-introduced and the antenna error, we set up an experiment consisting of a Trimble Zephyr antenna split into a GMU (ublox LEA-M8T single-frequency receiver), a Trimble BD930 receiver (geodetic dual-frequency), and an experimental low-cost dual-frequency receiver, developed by Saphyrion Sagl, capable of tracking L1 and L2C signals. An observation period from 9 to 17 February 2018 was selected. The location of the experimental setup and the SEID-reference stations used are depicted in Figure 8.

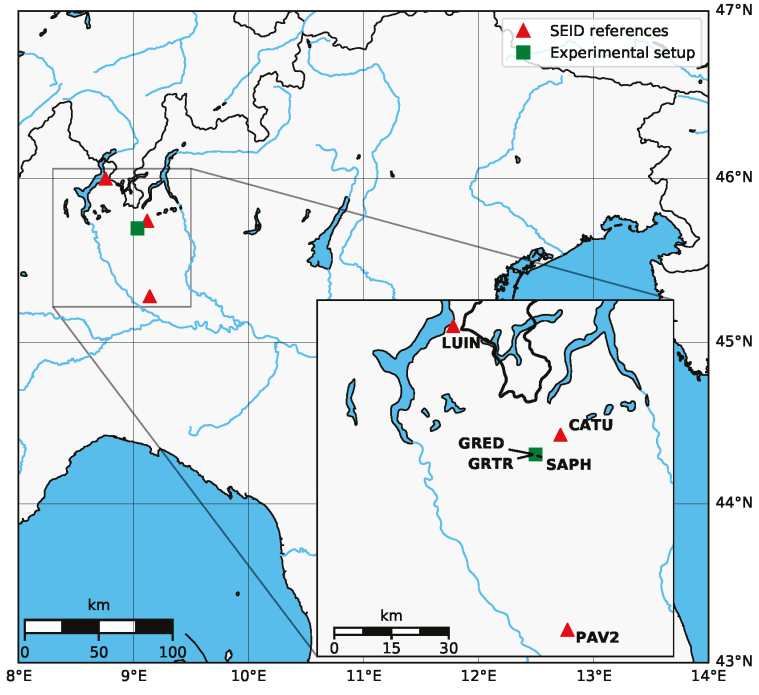
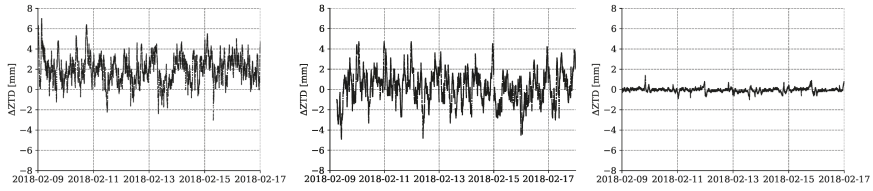


Figure 8. Positions of the Italy-experiment SEID reference stations constellation (red triangles) and experimental setup (green square). The sites GRTR, GRED, and SAPH use the same Trimble Zephyr 2 antenna (TRM55971.00). GRTR is the dual-frequency receiver reference (Trimble BD930 receiver), GRED the cost-efficient single-frequency device (u-blox), and SAPH a low-cost experimental dual-frequency receiver.

GRTR is the dual-frequency reference station, GRED is the u-blox LEA-M8T single-frequency receiver (same as GMU from the prior experiments), and SAPH is the experimental low-cost dual-frequency receiver. All units obtain the same antenna input signal from the geodetic GRTR antenna, through a GPS Source 4-way splitter. For our analysis, we utilized the SEID reference stations LUIN, PAV2, and CATU. They are part of the NetGEO [33] network and are located within 10–50 km from the experimental setup. After synthesizing a single-frequency dataset from GRTR (GRTR_SF), we applied the SEID algorithm to the single-frequency datasets (GRTR_SF and GRED).

On all datasets, the PPP was performed to obtain ZTD estimations at the locations. Contrary to the previous computations, this analysis utilized goGPS version 0.6.0, which is based on a joint least-squares adjustment. Figure 9 depicts the differences between selected time series.



(**a**) GRTR_SF - GRTR_DF (**b**) GRTR_DF - SAPH_DF (**c**) GRTR_SF - GRED_SF

Figure 9. ZTD differences between GRTR_SF and GRTR_DF (**a**), GRTR_DF and SAPH_DF (**b**), and GRTR_SF - GRED_SF (**c**). The GRTR_SF and GRED_SF estimations are based on the SEID algorithm, whilst GRTR_DF and SAPH_DF use PPP-only.

Figure 9a shows the differences between the synthesized dataset GRTR_SF and its dual-frequency counterpart GRTR_DF. A bias of 1.99 mm and an RMSE of 2.40 mm is present. The comparison between the geodetic grade GRTR_DF and the experimental cost-efficient dual-frequency receiver (Figure 9b) yields an RMSE of 1.58 mm. No significant bias is visible. Figure 9c shows the differences between the SEID-generated single-frequency ZTD estimations (GRTR_SF and GRED_SF). The comparison yields an RMSE of 0.20 mm.

4. Discussion

4.1. Inter-Comparison of Reference Datasets and Analysis of the Software-Related Error

For the IGS network stations, various ZTD reference datasets from different providers are available. In our experiment, we conducted an inter-comparison between 5 analysis center outputs. The results depict a very good agreement between these yielding standard deviations between 2 and 4 mm. The variation between the reference datasets corresponds well to the uncertainty of 4 mm of the IGS troposphere products compared to other independent measurement techniques like radiosondes or numerical weather models [28]. However, the goGPS-processed results highlight differences between the goGPS PPP estimations and the available ZTD reference datasets. In our experiment, the goGPS PPP version introduced a 4 mm ZTD bias to the dual-frequency PPP solution. A part of this error may be caused by using a 10 degree elevation cut-off angle instead of 7 degrees as used in the IGS reference dataset. The comparison between the PPP-only (DLF1_DF) and the IGS dataset demonstrated an RMSE of 7 mm. This effect is mainly caused by the software-introduced bias and exceeds the 4 mm error range by the IGS troposphere products. It suggests that there is still room for improvements to the processing engine. However, the software is continuously developing, and the results are constantly improved by implementing additional model functions. In the upcoming release, the processing will switch from using a Kalman filter to a joint least-squares adjustment. This version is currently in an experimental phase. However, initial results (see Section 3.4) from the least-squares version suggest a significant improvement of the ZTD estimation accuracy. Even though the original dual-frequency ZTD estimations are subject to be improved by utilizing another PPP analysis tool, driven by the observed ZTD variation the results suggest that the goGPS estimations can be used to representatively compare the ZTD estimations among each another. Using a different processing engine will not significantly influence the investigated variations.

4.2. SEID DPGA Experiment and Antenna Impact

The conducted SEID experiment using the DPGA constellation with distances between 50 and 80 km demonstrate the successful application to our domain. The RMSE of 4 mm between the DLF1_SF and its dual-frequency counterpart validate the results from prior studies by, e.g., Deng et al. [14]. DLF1_DF and DLF1_SF use the same geodetic-grade antenna and antenna phase center corrections. Since they share the same receiver, antenna, and configuration parameters, the error between these two is entirely introduced by the SEID network interpolation and demonstrates a reasonable accuracy. The error may be introduced by loss of lock, cycle slips, or the irregular nature of the ionosphere. The ZTD results from the newly placed single-frequency stations GMU and UBX are generally noisier. The GMU, being equipped with a Tallysman antenna and u-blox LEA-M8T receiver, demonstrated a slight performance decrease compared to the synthesized dataset. The GMU contains fewer data points than the UBX device, because, especially at the end of the year, the GMU suffered technical problems caused by the insufficient power supply generated by the solar panels during winter. GMU and UBX are equipped with a similar receiver type (respectively, u-blox LEA-M8T and u-blox NEO-M8T), are co-located and use the same ionosphere correction sources. However, the UBX site is affected by a systematic bias (about 6 mm), which is presumably caused by an increased multipath effect due to the lower quality antenna. This suggests that the antenna type and placement has a major influence on the precision. The authors of Pesyna et al. [34] report that patch antennas are characterized by a less effective signal reception and suffer more from multipath effects than geodetic-grade antennas. The antenna is a simple patch antenna and shielded by a metal plate on the ground. The horizontal placement on the rooftop may be a reason for this effect. The DPGA SEID experiment demonstrated a successful application of the SEID model to single-frequency receivers with standard deviations of the ZTD error between 4 and 5 mm. A decrease in the outer network station distance using an even denser network (e.g., the Japanese GEONET or existing commercial GNSS networks) could even further improve the results. However, not all regions on the Earth, such as Europe, North America, and Japan, provide such a dense network of dual-frequency GNSS receivers. Areas that will benefit most from a densification with cost-efficient single-frequency receivers are often characterized by a much coarser reference network.

4.3. SEID EUREF Experiment

Utilizing the EUREF stations with distances between 200 and 300 km as SEID references allowed us to experimentally investigate the impact of significantly longer station distances on the ZTD estimations. The comparison between DLF1_DF and DLF1_SF indicated an ionospheric error and yielded an RMSE of 10 mm. The overall character of the data shows more variation, which results in fewer outliers. The results also suggest that the GMU reaches a similar precision as the DLF1_SF dataset. However, the GMU and UBX stations have data only available from 1 September 2017 onwards, whereas the DLF1_SF provides data throughout the year. The higher ZTD values in summer are missing for GMU and UBX. A decrease in these stations' performance may be found if the full observation period is considered. The majority of the outliers at the experimental setup are presumably caused by hardware errors, inconsistencies during the pre-processing (e.g., undetected cycle slips during the SEID processing), re-initializations of the Kalman filter after missing epochs, and interpolation of the ionospheric component. The higher outliers are especially present at epochs with higher re-initialization rates that are caused by power-problems of the GMU. The GMU shows a slightly reduced removal of epochs (6.89%) compared to the DPGA case, which is presumably caused by less abrupt changes in ZTD estimations. This large reference network demonstrates that the limiting factor of the precision of the ZTD estimations is the ionospheric interpolation rather than the hardware quality. We expect that the GMU will provide comparable results with almost the same precision as the synthesized dataset since the noise introduced by the SEID model will dominate the estimations. This experiment shows that the SEID algorithm is able to provide ZTD estimations even over a SEID network with distances between 200 and 300 km to the single-frequency setup in a mid-latitude region.

However, the ZTD RMSE of 10–12 mm compared to the dual-frequency estimation is just barely acceptable for the post-processed water vapor estimation.

4.4. PWV Estimations

An additional error is introduced by the conversion from ZTD to PWV by utilizing temperature and pressure data. Without co-located temperature and air pressure measurements, data from weather models may be used (e.g., Jiang et al. [35]). In this study we use observations from a nearby weather station. Due to the small distance of about 5 km and almost negligible height differences, a linear interpolation of the temperature is sufficient. However, the heterogeneous character of surface temperature caused, e.g., by effects such as urban heat island will result in additional PWV uncertainty. Having achieved PWV RMSE values between 0.6–1.05 mm (DPGA case) and 1.61–1.86 mm (EUREF case) compared to the dual-frequency reference, it must be highlighted that the errors from GMU and UBX are especially subject to increase since the summer period is missing in the analysis. The results show that cost-efficient single-frequency devices may provide beneficial PWV estimations over areas with up to 300 km SEID reference station distances. Such an application may be especially interesting for regions with limited investment capabilities and areas that show a high PWV variability and no or only very few available PWV in-situ data. Extending existing tropical GNSS monitoring networks like the Continuously Operating Caribbean GPS Observational Network (COCONet) or the Amazon Dense Network [36] could potentially improve the monitoring and forecasting of deep tropical convection. However, especially the geomagnetic anomaly over South America and equatorial regions are characterized by a stronger and more turbulent ionosphere than mid-latitudes. Additional research is required to evaluate the applicability of the SEID algorithm for such regions.

4.5. Antenna Splitting

The conducted antenna splitting experiment allowed us to distinguish between the SEID-introduced and hardware errors. The results suggest that, overall, the experimental goGPS version (version 0.6.0) improved the ZTD estimation, and the introduced bias was reduced. Initialization times at the day-boundaries were smoothed by buffering for 3 h before and after each daily file. Figure 9a demonstrates that the bias in this case is introduced by the SEID implementation. The differences illustrate entirely the error that is introduced by the SEID implementation. Compared to the DPGA and EUREF experiments, the reduced variation of the ZTD estimations results partly from the version change and partly from the decreased SEID station distances. However, correcting the mean bias will improve the RMSE significantly. The cost-efficient dual-frequency SAPH_DF station yields an RMSE of 1.58 mm. The most notable characteristic of the SAPH_DF station is that it is only able to track up to 8 GPS satellites simultaneously since it is limited by the amount of concurrent receiving channels. The ZTD deviation is introduced by the decreased receiver quality and by a less optimal satellite geometry due to the fact that only up to 8 satellites on two frequencies may be observed. Difficult environmental conditions, e.g., increased multipaths, may even further decrease the quality. Since the same ionosphere correction, antenna data and pattern correction is used for GRTR_SF and GRED_SF, Figure 9c illustrates the receiver clock and phase ambiguity error. It clearly demonstrates that low-cost single-frequency receivers are capable to track satellite signals with almost identical precision to geodetic-grade receivers. This suggests that the precision of the single-frequency units for ZTD estimations is foremost dependent on the correct modeling of the ionosphere and the antenna type.

5. Conclusions

Higher spatial resolution of water vapor measurements is required for the observation of smaller scale convective events [37]. Prior studies, e.g., by Oigawa et al. [10], demonstrate a positive impact using a high resolution GNSS-based PWV network on the rainfall forecast accuracy. The objective of this work was to evaluate the feasibility of a densification of existing GNSS networks with cost-efficient single-frequency GNSS receivers. In this study, we analyzed one year of GNSS data. We first compared

the available ZTD reference datasets with each other at the experimental setup location. Secondly, we validated the ZTD estimations at our experimental setup using two different SEID reference networks and indicated the potential for PWV monitoring. Furthermore, we conducted an additional observation campaign to distinguish between the hardware and the SEID introduced errors.

We experimentally verified that the applied SEID model is able to successfully be used for water vapor estimations using cost-efficient single-frequency receivers. A clear precision degradation linked to the increasing outer network station distances was found. The SEID-processed synthesized dataset (DFL1_SF) demonstrated a ZTD RMSE of 3.93 and 10.32 mm for station distances of up to 80 and 300 km, respectively. The cost-efficient unit (GMU) indicates comparable results with RMSE values of 5.55 and 10.20 mm. The precision further decreased by the utilization of a simple patch antenna (station UBX) to RMSE values of 7.10 and 12.09 mm. A clear bias of −4.96 mm, presumably caused by the antenna type or an increased multipath effect, was observed for this receiver. Our analysis also highlighted a systematic ZTD bias of about 4 mm caused by the used goGPS version 0.4.3 compared to existing ZTD reference datasets. However, apart from the bias, the precision of the PPP ZTD estimations were found to be reliable.

The additional antenna splitting experiment conducted in Italy using the newest goGPS version 0.6.0 demonstrated a significant improvement to the ZTD estimation. However, the decreased SEID reference station distance also positively affected the ZTD errors. A ZTD RMSE of 2.40 mm was found for the synthesized single-frequency dataset (GRTR_SF) compared to its dual-frequency counterpart. A cost-efficient dual-frequency receiver that tracks only a limited number of satellites showed an RMSE of 1.58 mm in ZTD. Using only a few satellites, the satellite geometry deteriorates and the ZTD estimation may be negatively influenced. The experiment demonstrated that, using the SEID model, cost-efficient single-frequency receivers are able to provide ZTD estimations with almost negligible differences.

The study demonstrated that cost-efficient single-frequency receivers can serve as a promising complement to the presently available GNSS networks in mid-latitude regions. Cost-efficient single-frequency stations like the GMU used in this analysis are in operational use since the end of 2015. Historical data about the longevity for these recently designed units is not yet available. However, fulfilling certified industrial standards for protection against sand, dust, and water provide the best conditions for a long durability of the stations. In the framework of the TWIGA project [38], a network of such GMUs will be deployed in equatorial regions in Africa. The aim is to investigate the durability in more challenging environments and how the SEID model will perform in the equatorial region with more turbulent ionospheric conditions. The conducted experiments also demonstrate that the ZTD precision from single-frequency receivers is foremost dependent on the modeling of the ionosphere and the antenna type than on the internal hardware. However, the application of other ionospheric correction models is not discussed in detail in this analysis and a comparison to different ionospheric models must be considered for future studies.

Further work is required to utilize the multi-GNSS ability of the receivers for the SEID approach. Additional satellites combined with a denser network of GNSS receivers may be useful for GNSS-based 3D tomography. Especially GNSS tomography can be improved by using a multi-GNSS approach [39]. It must be highlighted that all shown results are post-processed PPP solutions, available with a temporal delay of at least two weeks. For an effective weather forecasting, the data needs to be available in near-real time (NRT). The use of multi-GNSS PPP utilizing NRT data from the IGS has been studied by Lu et al. [40]. Compared to the GPS-only mode, they were able to improve the accuracy of ZTD estimations by up to 22.2%.

Densified PWV networks may be used to improve the long-term accuracy of weather models, for climate analysis, or to improve model calibrations. They can be beneficial when studying particular weather events at particular sites or on a regional scale. In the framework of BRIGAID, the Rotterdam study area (The Netherlands) allows us to make use of additional existing dual-frequency stations within the city borders. The decreased reference station distances are expected to improve ionospheric modeling and to result in better PWV estimations. The target of our future work is to study the spatial

variability of single-frequency-based PWV estimations and to analyze its impact on the monitoring and forecasting of convective events in that area. For this purpose, we installed a follow-up experimental single-frequency network consisting of four additional receivers.

Author Contributions: All authors contributed to the design of the research plan. A.K. set up the experiment in Delft, processed the datasets, performed post-processing analyses, and wrote the paper. H.v.d.M. maintained the DPGA data collection. E.R. conducted the antenna splitting experiment in Italy. M.-c.t.V. and N.v.d.G. contributed to the research design, data analysis and final editing of the manuscript.

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700699.

Acknowledgments: We like to thank Peter F. de Bakker for assistance with the configuration and installation of the u-blox receiver in Delft. Besides all goGPS contributors, the authors would like to express their gratitude to Andrea Gatti and Giulio Tagliaferro for their insightful comments and maintaining the goGPS software.

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

Abbreviations

The following abbreviations are used in this manuscript:

AC	Analysis Center
AWS	Automatic Weather Station
BKG	Federal Agency for Cartography and Geodesy
CDDIS	Crustal Dynamics Data Information System
COCONet	Continuously Operating Carribbean GPS Observational Network
DLF1_DF	IGS station DLF1 (dual-frequency)
DLF1_SF	IGS station DLF1 (single-frequency)
DPGA	Dutch Permanent GNSS Array
EGM08	Earth Gravity Model 2008
E-GVAP	EUMETNET GNSS water vapor Programme
EPN	EUREF Permanent Network
GFZ	German Research Centre for Geosciences
GMF	Global Mapping Function
GMU	GeoGuard Monitoring Unit
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPT	Global Pressure/Temperature
IGS	International GNSS Service
IP	Ingress Protection
IPP	Ionospheric Pierce Point
IWV	Integrated Water Vapor
KNMI	Royal Netherlands Meteorological Institute
LPT	Federal Office of Topography
NGL	Nevada Geodetic Laboratory
NRT	Near-Real Time
PPP	Precise Point Positioning
PWV	Precipitable Water Vapor
RF	Radio Frequency
RINEX	Receiver Independent Exchange Format
RMSE	Root Mean Square Error
ROB	Royal Observatory of Belgium
RTK	Real-Time Kinematic
SEID	Satellite-specific and Epoch-differenced Ionospheric Delay
STD	Slant Total Delay
teqc	translation, editing and quality check
UBX	u-blox NEO-M8T evaluation toolkit
ZHD	Zenith Hydrostatic Delay
ZTD	Zenith Tropospheric Delay
ZWD	Zenith Wet Delay

References

1. Kiehl, J.T.; Trenberth, K.E. Earth's Annual Global Mean Energy Budget. *Bull. Am. Meteorol. Soc.* **1997**, *78*, 197–208. [CrossRef]
2. Andrews, D. *An Introduction to Atmospheric Physics*; International Geophysics Series; Cambridge University Press: Cambridge, UK, 2000.
3. Seko, H.; Nakamura, H.; Shoji, Y.; Iwabuchi, T. The meso-y scale water vapor distribution associated with a thunderstorm calculated from a dense network of GPS receivers. *JMSJ* **2004**, *82*, 569–586. [CrossRef]
4. Cess, R.D.; Potter, G.L.; Blanchet, J.P.; Boer, G.J.; Del Genio, A.D.; Déqué, M.; Dymnikov, V.; Galin, V.; Gates, W.L.; Ghan, S.J.; et al. Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J. Geophys. Res. Atmos.* **1990**, *95*, 16601–16615. [CrossRef]
5. Ning, T.; Wang, J.; Elgered, G.; Dick, G.; Wickert, J.; Bradke, M.; Sommer, M.; Querel, R.; Smale, D. The uncertainty of the atmospheric integrated water vapour estimated from GNSS observations. *Atmos. Meas. Tech.* **2016**, *9*, 79. [CrossRef]
6. Kuo, Y.H.; Zou, X.; Guo, Y.R. Variational assimilation of precipitable water using a nonhydrostatic mesoscale adjoint model. Part I: Moisture retrieval and sensitivity experiments. *Mon. Weather Rev.* **1996**, *124*, 122–147. [CrossRef]
7. Sato, K.; Realini, E.; Tsuda, T.; Oigawa, M.; Iwaki, Y.; Shoji, Y.; Seko, H. A high-resolution, precipitable water vapor monitoring system using a dense network of GNSS receivers. *J. Dis. Res.* **2013**, *8*, 37–47. [CrossRef]
8. Realini, E.; Sato, K.; Tsuda, T.; Oigawa, M.; Iwaki, Y.; Shoji, Y.; Seko, H. Local-scale Precipitable Water Vapor Retrieval from High-Elevation Slant Tropospheric Delays Using A Dense Network of GNSS Receivers. In *IAG 150 Years. International Association of Geodesy Symposia*; Rizos, C., Willis, P., Eds.; Springer: Cham, Switzerland, 2015; Volume 143, pp. 485–490.
9. GNSS Earth Observation Network System. Available online: http://datahouse1.gsi.go.jp/terras/terras_english.html (accessed on 12 September 2018).
10. Oigawa, M.; Tsuda, T.; Seko, H.; Shoji, Y.; Realini, E. Data assimilation experiment of precipitable water vapor observed by a hyper-dense GNSS receiver network using a nested NHM-LETKF system. *Earth Planets Space* **2018**, *70*, 74. [CrossRef]
11. Baldysz, Z.; Nykiel, G.; Figurski, M.; Araszkiewicz, A. Assessment of the Impact of GNSS Processing Strategies on the Long-Term Parameters of 20 Years IWV Time Series. *Remote Sens.* **2018**, *10*, 496. [CrossRef]
12. 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]
13. 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]
14. Deng, Z.; Bender, M.; Zus, F.; Ge, M.; Dick, G.; Ramatschi, M.; Wickert, J.; Lhnert, U.; Schön, S. Validation of tropospheric slant path delays derived from single and dual frequency GPS receivers. *Radio Sci.* **2011**, *46*, 1–11. [CrossRef]
15. GeoGuard Earth Monitoring Services. Available online: http://www.geoguard.eu/ (accessed on 16 May 2018).
16. Thayer, G. An improved equation for the radio refractive index of air. *Radio Sci.* **1974**, *9*, 803–807. [CrossRef]
17. Saastamoinen, J. Contributions to the theory of atmospheric refraction. *Bull. Geod.* **1972**, *46*, 279–298. [CrossRef]
18. 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]
19. Baltink, H.K.; Van Der Marel, H.; van der Hoeven, A.G. Integrated atmospheric water vapor estimates from a regional GPS network. *J. Geophys. Res. Atmos.* **2002**, *107*. [CrossRef]
20. Bevis, M. GPS meteorology: Mapping zenith wet delays onto precipitable water. *J. Appl. Meteorol.* **1994**, *33*, 379–386. [CrossRef]
21. Pavlis, N.K.; Holmes, S.A.; Kenyon, S.C.; Factor, J.K. An earth gravitational model to degree 2160: EGM2008. *EGU Gener. Assem.* **2008**, *10*, 13–18.
22. Berberan-Santos, M.N.; Bodunov, E.N.; Pogliani, L. On the barometric formula. *Am. J. Phys.* **1997**, *65*, 404–412. [CrossRef]

23. Schaer, S. Mapping and predicting the Earth's ionosphere using the Global Positioning System. *Geod. Geophys. Arb. Schweiz* **1999**, *59*.
24. Araujo-Pradere, E.; Fuller-Rowell, T.; Spencer, P.; Minter, C. Differential validation of the US-TEC model. *Radio Sci.* **2007**, *42*, RS3016. [CrossRef]
25. Estey, L.H.; Meertens, C.M. TEQC: The Multi-Purpose Toolkit for GPS/GLONASS Data. *GPS Solut.* **1999**, *3*, 42–49. [CrossRef]
26. 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]
27. Lyard, F.; Lefevre, F.; Letellier, T.; Francis, O. Modelling the global ocean tides: Modern insights from FES2004. *Ocean. Dynam.* **2006**, *56*, 394–415. [CrossRef]
28. 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]
29. 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]
30. Blewitt, G. FTP Archive of NGL Troposphere Delay Products. Available online: ftp://gneiss.nbmg.unr.edu/trop (accessed on 16 May 2018).
31. EUREF Permanent GNSS Network. Available online: http://epncb.eu/_productsservices/troposphere/ (accessed on 16 May 2018).
32. de Haan, S. *National/Regional Operational Procedures of GPS Water Vapour Networks and Agreed International Procedures*; Rep WMO/TD-No. 1340; World Meteorological Organization: Geneva, Switzerland, 2006.
33. Topcon Positioning Italy. Available online: http://www.netgeo.it/ (accessed on 28 May 2018).
34. Pesyna, K.M.J.; Robert, W.H.J.; Humphreys, T.E. Centimeter positioning with a smartphone-quality GNSS antenna. In Proceedings of the ION GNSS+ Meeting, Tampa, FL, USA, 8–12 September 2014.
35. Jiang, P.; Ye, S.; Chen, D.; Liu, Y.; Xia, P. Retrieving precipitable water vapor data using GPS zenith delays and global reanalysis data in China. *Remote Sens.* **2016**, *8*, 389. [CrossRef]
36. Adams, D.K.; Fernandes, R.M.; Kursinski, E.R.; Maia, J.M.; Sapucci, L.F.; Machado, L.A.; Vitorello, I.; Monico, J.F.G.; Holub, K.L.; Gutman, S.I.; et al. A dense GNSS meteorological network for observing deep convection in the Amazon. *Atmos. Sci. Lett.* **2011**, *12*, 207–212. [CrossRef]
37. 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]
38. Transforming Weather Water Data into Value-Added Information services for Sustainable Growth in Africa. Available online: http://twiga-h2020.eu/ (accessed on 12 September 2018).
39. Dong, Z.; Jin, S. 3-D Water Vapor Tomography in Wuhan from GPS, BDS and GLONASS Observations. *Remote Sens.* **2018**, *10*, 62. [CrossRef]
40. Lu, C.; Chen, X.; Liu, G.; Dick, G.; Wickert, J.; Jiang, X.; Zheng, K.; Schuh, H. Real-Time Tropospheric Delays Retrieved from Multi-GNSS Observations and IGS Real-Time Product Streams. *Remote Sens.* **2017**, *9*, 1317. [CrossRef]

remote sensing

MDPI

Article

Comparison and Bias Correction of TMPA Precipitation Products over the Lower Part of Red–Thai Binh River Basin of Vietnam

Hung Manh Le [1,2,*], Jessica R. P. Sutton [3], Duong Du Bui [2], John D. Bolten [4] and Venkataraman Lakshmi [1]

[1] School of Earth, Ocean and the Environment, University of South Carolina, Columbia, SC 29208, USA; vlakshmi@geol.sc.edu
[2] National Center for Water Resources Planning and Investigation, Ministry of Natural Resources and Environmental, Hanoi 100000, Vietnam; duongdubui@gmail.com
[3] Berry College, Mount Berry, GA 30149, USA; jsutton@berry.edu
[4] Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; john.bolten@nasa.gov
* Correspondence: hmle@email.sc.edu; Tel.: +1-702-861-4265

Received: 5 September 2018; Accepted: 24 September 2018; Published: 1 October 2018

Abstract: As the limitation of rainfall collection by ground measurement has been widely recognized, satellite-based rainfall estimate is a promising high-resolution alternative in both time and space. This study is aimed at exploring the capacity of the satellite-based rainfall product Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), including 3B42V7 research data and its real-time 3B42RT data, by comparing them against data from 29 ground observation stations over the lower part of the Red–Thai Binh River Basin from March 2000 to December 2016. Various statistical metrics were applied to evaluate the TMPA products. The results showed that both 3B42V7 and 3B42RT had weak relationships with daily observations, but 3B42V7 data had strong agreement on the monthly scale compared to 3B42RT. Seasonal analysis showed that 3B42V7 and 3B42RT underestimated rainfall during the dry season and overestimated rainfall during the wet season, with high bias observed for 3B42RT. In addition, detection metrics demonstrated that TMPA products could detect rainfall events in the wet season much better than in the dry season. When rainfall intensity was analyzed, both 3B42V7 and 3B42RT overestimated the no rainfall event during the dry season but underestimated these events during the wet season. Finally, based on the moderate correlation between climatology–topography characteristics and correction factors of linear-scaling (LS) approach, a set of multiple linear models was developed to reduce the error between TMPA products and the observations. The results showed that climatology–topography-based linear-scaling approach (CTLS) significantly reduced the percentage bias (PBIAS) score and moderately improved the Nash–Sutcliffe efficiency (NSE) score. The finding of this paper gives an overview of the capacity of TMPA products in the lower part of the Red–Thai Binh River Basin regarding water resource applications and provides a simple bias correction that can be used to improve the correctness of TMPA products.

Keywords: Red–Thai Binh River Basin; TMPA 3B42V7; TMPA 3B42RT; rainfall; bias correction; linear-scaling approach; climatology; topography

1. Introduction

Precipitation is the most crucial input variable enforced in water prediction models. Reliable precipitation is required for model calibration, forecast, and simulation [1–3]. Gauge observation is

the primary collection approach to obtain precipitation information [4]. However, gauge network is often sparse and nonexistent in many parts of the globe [5,6]. Moreover, it is often challenging to obtain gauge data, especially in developing countries and transboundary rivers, due to technical and administrative reasons [7–9]. In addition, gauge observations only provide point measurements of precipitation and cannot capture the full spatial variability. Space-based precipitation estimations, therefore, have great potential application to enhance the capacity of measuring this vital water cycle component [10,11].

Several satellite-derived datasets have been used in previous studies, such as the Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) [12], the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [13], the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) [14], and National Oceanic and Atmospheric Administration/Climate Prediction Centre (NOAA/CPC) morphing technique (CMORPH) [15] products. Among them, TPMA—the first space-borne product of the Earth Science Mission aimed at studying tropical and subtropical rainfall—has performed well in a wide range of applications, such as hydrological modeling [16–18], drought monitoring [19,20], and agronomy [21,22]. TMPA products have also been evaluated as having better performance than other satellite-based rainfall products. For example, the TMPA 3B42V7 data is generally a better input in a distributed hydrological model compared to CMORPH and TMPA 3B42RT (real time) for multiple hydrological purposes, including annual water budgeting, monthly and daily streamflow simulation, and extreme flood modeling [23]. Similarity, Tong et al. [24] showed that 3B42V7 was a better driving force of hydrological model for both monthly and daily streamflow simulation over the Tibetan Plateau compared to CMORPH, PERSIANN, and 3B42RT. Moazami et al. used six statistical indices and contingency table to evaluate 3B42V7, concluding it was a better estimation of daily precipitation than PERSIAN and 3B42RT over Iran [25]. Simons et al. [26] identified that monthly TMPA 3B43 rainfall product was the most suitable satellite dataset compared to CHIRPS and CMORPH over the Red River Basin of Vietnam.

Differences between TMPA products and rain gauge observation analysis have been a cause of concern recently. Zad et al. [27] pointed out that 3B42V7 tended to overestimate rainfall measurement by approximately 26.95% at Pahang River Basin of Malaysia and that 3B42V7 was likely to have a high accuracy of detecting rainfall events at high-altitude and mid-altitude areas compared to low-altitude regions. Kneis et al. [28] analyzed that 3B42V7 and 3B42RT datasets were moderately correlated with their gauged-based counterpart at sub-basin level (4000 to 16,000 km^2) at the lower Mahanadi River Basin of India but that the 3B42V7 and 3B42RT data often do not reflect gauge observation at high-intensity level (>80 mm/day). The TMPA product is also likely to perform better on a monthly scale when compared to the ground data. Curtarelli et al. [29] found that monthly 3B43 dataset had a great consistency (correlation coefficient >0.97) with ground observation data over the Itumbiara Reservoir drainage area in Central Brazil but that 3B43 tended to overestimate rainfall by 1.24%. Comparing monthly 3B43 dataset with 56 observations in Yangtze River Delta, Cao et al. [30] also showed an inclination of 3B43 to overestimate monthly rainfall, with the bias ranging between −10% and 10% most of the study area; its correlation coefficient with observation was found to peak in March (0.96) and reach bottom in August (0.79). Although the TRMM satellite has not been operated since 2014, TMPA products are still being generated regardless [31].

Following the highly successful TMPA, the Global Precipitation Measurement (GPM) mission was developed to continuously increase precipitation estimation over most of the globe [32]. A range of studies in many regions have demonstrated that GPM outperforms TMPA by having a better spatial resolution, coverage area, and lower systematic bias error [33–35]. However, GPM has only been available for a short time (since 2014), while TMPA products date back to January 1998. In addition, GPM is just a slight improvement over TMPA products [36]. Huffman et al. [32] aim to extend the GPM data to the same length as the longest TMPA data. Therefore, assessments on TMPA products

are of paramount importance to gain insights into their performance at various regions so that their algorithms can be improved and the next generation GPMs can be developed.

While there is a clear advantage of having a high temporal and spatial resolution using TMPA products, extra work is required because bias correction needs to be performed prior to application of any TMPA products in environmental, water resources, and ecological studies [27]. Climatology and topography are likely factors to induce errors in remote sensing retrievals [37]. Consequently, their effects on the quality of TMPA products are inevitable. Based on the moderate inverse linear relationship between the monthly 3B43 bias and elevation, Hashemi et al. [38] developed a linear model between 3B43 bias and elevation, especially for stations that have elevations above 1500 m above mean sea level in the U.S. The corrected monthly 3B43 product showed a significant improvement in the high elevation area. Thus, the empirical bias correction model using climatology and topography seems to be a potential investigation direction, although relatively little research has been conducted so far.

In Vietnam, ground observations provide poor spatial and temporal measurement of rainfall due to the lack of a dense network for rain gauge measurement. The average rain gauge network in Vietnam is around 400 km^2 per rain gauge, which is below the World Meteorological Organization standard (area per rainfall station of 100–250 km^2 for mountainous areas; area per rainfall station of 600–900 km^2 for lowland areas) [39]. Moreover, the rain gauge distribution in Vietnam is uneven, with insufficient gauged stations at high elevation areas. According to the Vietnam Meteorological and Hydrological Administration, most rain gauge stations (75%) are concentrated at low elevation areas (<200 m), which only cover half of Vietnam's land [40]. With these perspectives, satellite-based precipitation is an indispensable alternative source of precipitation data for Vietnam. Preliminary studies on satellite-based precipitation products in the country have been conducted recently. However, these studies either focused on monthly rainfall [26,41] or used directly satellite-derived rainfall without bias correction analysis [42]. Therefore, further research on satellite-based precipitation products is still of fundamental importance for the country.

This study selected the Red–Thai Binh River Basin—one of the largest river systems in Vietnam—as a case study. Although it plays an essential role in Vietnam's economic and social development, many parts of this basin do not have rainfall monitoring from ground, causing difficulties for basin rainfall estimation and water resources management. The first objective of this study was to compare the TMPA products 3B42V7 and 3B42RT with ground observation data over Red–Thai Binh River Basin in various aspects, such as calculating error statistics on a daily scale, monthly scale, dry and wet seasons, detecting rainfall events ability, and evaluating rainfall intensity. The second objective was to develop a linear-scaling bias correction model using climate–topography indices for both 3B42V7 and 3B42RT datasets. The results of the assessment and bias correction of TMPA precipitation products could help in supporting its potential application in hydrological modeling and drought monitoring in the studied region.

2. Materials

2.1. Study Area

The Red–Thai Binh River Basin is a transboundary river that flows through three countries—Vietnam, China, and Laos—with a total area of 169,000 km^2 (Figure 1). The area of this in Vietnam is 88,680 km^2, which makes up 51.3% of the total area. In this study, due to the lack of observation data, description of water resource characteristics and evaluation results of TMPA 3B42V7 and TMPA 3B42RT data only focused on the Vietnamese part of the basin. There are two primary river systems in the Red–Thai Binh River. The Red River system originates in China and flows into Vietnam through three main tributaries—Da, Lo, and the Thao River—while the Thai Binh River system is entirely located in Vietnam. The Red–Thai Binh River belongs to a tropical climate with two distinct seasons: the wet season and the dry season. The total annual rainfall is approximately 1700 mm, with high rainfall amounts (>2000 mm) observed in the mountainous areas between the

Vietnam and China border. The annual total flow of the Red–Thai Binh River is 131.4 billion m^3—the Chinese territory part generates 48.3 billion m^3, while the rest 83.1 billion m^3 is generated in the Vietnamese side [43]. As the second largest river system in Vietnam, the Red–Thai Binh River is home to 29.1 million Vietnamese (2015 figure), making up for 22.6% of Vietnam's GDP (2010 figure) (General Statistics of Vietnam) [44].

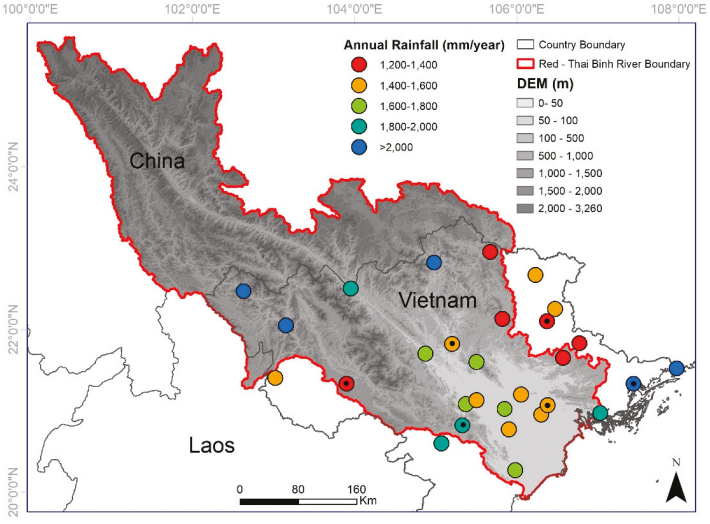


Figure 1. Overview of Red–Thai Binh River Basin. The stations with black dots at the middle were used for calibration climatology–topography-based linear-scaling approach.

2.2. Data

2.2.1. Observation Data

Rainfall measurements from a total of 29 daily rainfall stations (March 2000 to December 2016) within or neighboring the basin were collected from the Vietnam Meteorological and Hydrological Administration. The distribution of rainfall stations is presented in Figure 1, and their characteristics can be found in Table 1. The stations were selected due to their reliable data and low missing values (5–10%).

Table 1. Rainfall station descriptions for the ground observation stations over Red–Thai Binh River Basin (March 2000–December 2016).

No.	Station Name	Long. (°)	Lat. (°)	Elev. (m)	Annual Rainfall (AR) (mm/year)	Standard Deviation of Rainfall (SDR) (mm/day)	No. of Rain Days (NRD) (day)
1	Baccan	105.82	22.13	241	1389	11.29	250
2	Bacninh	106.05	21.20	8	1537	13.27	248
3	Baolac	105.67	22.95	348	1201	9.91	263
4	Caobang	106.23	22.67	244	1417	11.5	251
5	Dienbien	103.02	21.40	487	1535	11.75	248
6	Hagiang	104.98	22.82	117	2333	17.15	222
7	Bavi	105.37	21.08	535	1791	14.61	234
8	Lang	105.85	21.02	17	1686	14.5	246
9	Phuxuyen	105.90	20.77	9	1516	13.16	272
10	Sontay	105.50	21.13	14	1600	13.27	244
11	Chilinh	106.38	21.07	1	1489	12.51	250
12	Haiduong	106.30	20.95	3	1530	13.6	249

Table 1. *Cont.*

No.	Station Name	Long. (°)	Lat. (°)	Elev. (m)	Annual Rainfall (AR) (mm/year)	Standard Deviation of Rainfall (SDR) (mm/day)	No. of Rain Days (NRD) (day)
13	Hoabinh	105.33	20.82	48	1861	14.69	239
14	Maichau	105.07	20.60	579	1859	18.79	251
15	Muongte	102.63	22.47	354	2433	17.1	229
16	Tamduong	103.15	22.05	303	2333	14.49	216
17	Chilang	106.57	21.65	124	1324	11.9	267
18	Langson	106.77	21.83	263	1315	11.55	253
19	Thatkhe	106.47	22.25	157	1484	12.33	243
20	Vanmich	106.37	22.10	238	1341	11.33	240
21	Laocai	103.95	22.50	152	1810	14.11	229
22	Ninhbinh	105.98	20.27	3	1725	15.16	242
23	Baichay	107.03	20.97	59	1898	17.93	246
24	Mongcai	107.97	21.52	7	2735	24.5	230
25	Tienyen	107.44	21.33	16	2139	19.04	231
26	Sonla	103.90	21.33	709	1364	10.96	252
27	Thainguyen	105.50	21.60	784	1760	15.02	238
28	Tuyenquang	105.20	21.82	29	1575	14.02	242
29	Yenbai	104.87	21.70	41	1796	14.97	222

In Vietnam, daily ground rainfall data is often collected twice per day at 7.00 a.m. UTC + 7 and 7.00 p.m. UTC + 7, and the daily accumulation is calculated as accumulated rainfall from 7.00 a.m. UTC + 7 to the same time next day [45]. Figure 2 shows monthly rainfall distribution over Red–Thai Binh River Basin from gauge observation data. Wet season (May–October) has a high amount of rainfall, accounting for 85–90% of total annual rainfall. Very high amounts of rainfall are often observed during June, July, and August. During these periods, tropical storms often occur, with the accumulated rainfall reaching 200–600 mm within several days [44]. During the dry season (November–April), the total amount of rainfall only accounts for 10–15% of total annual rainfall.

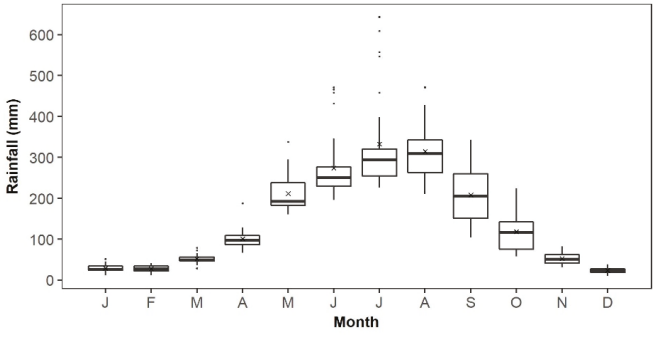


Figure 2. Monthly rainfall distribution over Red–Thai Binh River Basin (March 2000–December 2016). Cross symbol indicates average monthly rainfall.

2.2.2. TMPA Products

The TRMM is a low Earth orbits (LEO) satellite with sensors used to analyze and understand the characteristics of precipitation. The satellite is equipped with various instruments, such as Precipitation Radar (PR), TRMM Microwave Imager (TMI), Visible and Infrared Scanner (VIRS), and Lightning Imaging Sensor (LIS) [12]. The spatial coverage of TRMM is mainly in tropical and subtropical zones (50°S to 50°N) from an altitude of 400 km. The TMPA products used in this study were TMPA 3B42V7 and its real-time version TMPA 3B42RT at 0.25° spatial resolution. Detailed description of 3B42V7

can be found in Reference [12] and that of 3B42RT can be found in Reference [46]. The 3B42V7 dataset ranges from January 1998 to present, while the 3B42RT product ranges from March 2000 to present. However, for comparison purpose, a consistent data length was required and data was therefore collected from March 2000 to December 2016 for both TMPA 3B42V7 and TMPA 3B42RT. Both products were downloaded through NASA Goddard Space Flight Center (https://pmm.nasa.gov/data-access/downloads/trmm/). In order to match the satellite rainfall products with the daily precipitation gauge data, the 3-hourly 3B42 products were accumulated to daily values at 0.00 UTC (equivalent to 7.00 a.m. UTC + 7).

3. Method

The comparison of TMPA 3B42V7 and TMPA 3B42RT precipitation against the ground observation data involved the extraction of data time series of TMPA products at the corresponding locations of the 29 meteorological stations. As one TMPA pixel contained one rainfall station, a total of 29 TMPA pixels were extracted to form the time series corresponding to the ground observation data.

3.1. Error Metric Assessment

To compare rainfall values between TMPA products and ground observation data, widely accepted error metrics—correlation coefficient (CC), Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE), and percent bias (PBIAS)—were used [47,48]. The formulas for the statistical metrics are presented as follows:

$$\mathrm{CC} = \frac{\sum_{i=1}^{N}\left(OBS_i - \overline{OBS}\right)\left(TMPA_i - \overline{TMPA}\right)}{\sqrt{\sum_{i=1}^{N}\left(OBS_i - \overline{OBS}\right)^2 \sum_{i=1}^{N}\left(TMPA_i - \overline{TMPA}\right)^2}} \quad (1)$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(TMPA_i - OBS_i\right)^2}{\sum_{i=1}^{N}\left(OBS_i - \overline{OBS}\right)^2} \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(TMPA_i - OBS_i\right)^2} \quad (3)$$

$$\mathrm{PBIAS} = 100\,\frac{\sum_{i=1}^{N}\left(TMPA_i - OBS_i\right)}{\sum_{i=1}^{N} OBS_i} \quad (4)$$

where N is the total of samples, OBS_i and $TMPA_i$ represent the rainfall values for the ground observation data and the TMPA data, respectively, and $\overline{OBS}$ and $\overline{TMPA}$ represent the mean of the corresponding variables. CC ranges from -1 to 1, with strong positive correlation when the CC value is closer to 1 and strong negative correlation when the CC value is closer to -1. NSE varies between $-\infty$ to 1, indicating how well the plot of satellite product values and ground values fit the 1:1 line. A NSE value closer to 1 indicates a more perfect match between satellite product and ground data. RMSE is unit-based and would shed further light on the accuracy of the TMPA products. PBIAS measures the average tendency of the satellite values to be larger or smaller than the corresponding ground observations.

3.2. Detection Metric Assessment

The probability of detection (POD), false alarm ratio (FAR), the probability of false detection (POFD), and critical success index (CSI) were used to compare the occurrence and nonoccurrence of rainfall events between TMPA products and ground data [27,33,49]. The POD was the ratio of the total number of rainfall events correctly detected by the TMPA products to the total number of actual rainfall events. The FAR evaluated the ratio of the number of rainfall falsely detected by the TMPA products to the total rainfall events estimated by the TMPA products. The POFD was a fraction of false

events detected by the TMPA products versus the correct observations of no rainfall events by the TMPA products. The CSI, which is a function of POD and FAR, was the most accurate detection metric. The rainfall day threshold was set as 0.6 mm/day, which was defined as a threshold between no rainfall event and low rainfall event within 24 h based on long-term rainfall analysis over Vietnam [50]. These detection metrics can be computed as follows:

$$\text{POD} = \frac{\text{Hits}}{\text{Hits} + \text{Misses}} \tag{5}$$

$$\text{FAR} = \frac{\text{False Alarms}}{\text{Hits} + \text{False Alarms}} \tag{6}$$

$$\text{POFD} = \frac{\text{False Alarms}}{\text{False Alarms} + \text{Correct Rejections}} \tag{7}$$

$$\text{CSI} = \frac{\text{Hits}}{\text{Hits} + \text{False Alarms} + \text{Misses}} \tag{8}$$

The Hit, Miss, False Alarm, and Correct Rejection are presented in a contingency table in Table 2. The perfect scores of the POD and CSI are 1, while the perfect scores of the POFD and FAR are 0.

Table 2. Contingency table to measure the correspondence between ground observation data and Tropical Rainfall Measurement Mission Multi-satellite Precipitation Analysis (TMPA) product concerning the threshold intensity of 0.6 mm/day of a point-to-point event [51].

		Ground Observation	
		Yes	**No**
TMPA Product	Yes	Hit	False Alarm
	No	Miss	Correct Rejection

3.3. Rainfall Intensity Evaluation

To evaluate the rainfall intensity, we used probability density function (PDF) to classify the daily rainfall amounts into six categories based on Vietnam's regulation on rainfall classification [50]: (1) 0 to 0.6 mm; (2) 0.6 to 6 mm; (3) 6 to 16 mm; (4) 16 to 50 mm; (5) 50 to 100 mm; (6) >100 mm. The PDF analysis has been previously applied for comparing satellite rainfall products and ground data in several studies [33,36].

3.4. Climate–Topography-Based Linear-Scaling (CTLS) Bias Correction Approach

The linear-scaling (LS) approach [52,53] was based on monthly correction factor, which was the ratio between long-term monthly mean data for ground observation and TMPA.

$$CF_m = \frac{\overline{OBS}_m}{\overline{TMPA}_m} \tag{9}$$

$$TMPA_{i,m}^{corrected} = TMPA_{i,m} \times CF_m \tag{10}$$

where CF_m is the monthly mean change factor at month m, $\overline{OBS}_m$ and $\overline{TMPA}_m$ represent the mean of ground observation and TMPA data at month m, respectively. $TMPA_{i,m}^{corrected}$ and $TMPA_{i,m}$ are the corrected TMPA data and original TMPA data at day i of month m, respectively. In this study, we developed a set of multiple linear models that predicted correction factors CF_m from climatology–topography characteristics. We acquired station information as longitude (LONG), latitude (LAT), elevation (ELEV), annual rainfall (AR), standard deviation of rainfall (SDR), and the number of rainfall day (NRD). The CF_m can be computed as follows:

$$CF_m = \alpha_{0m} + \alpha_{1m}LONG + \alpha_{2m}LAT + \alpha_{3m}ELEV + \alpha_{4m}AR + \alpha_{5m}SDR + \alpha_{6m}NRD \tag{11}$$

where α_{0m}, α_{1m}, α_{2m}, α_{3m}, α_{4m}, α_{5m}, α_{6m} are regression coefficients corresponding to correction factor at month m. In other words, we developed a set of 12 multiple linear models to estimate correction factors from climatology–topography data. In order to select the most suitable candidates for each multiple linear model, we analyzed the relationship between the correction factor and climatology–topography for a single month and selected the significant correlation candidates. We used 23 meteorological stations (80%) to develop the abovementioned multiple linear models and six meteorological stations (20%) to verify the models.

4. Results and Discussion

4.1. Comparison between TMPA Products and Ground Observation Data

4.1.1. Daily and Monthly Scale Assessment

Table 3 presents the TMPA 3B42V7 and TMPA 3B42RT data in daily scale and monthly scale performance over the Red–Thai Binh River compared to the ground observation stations for 17 years (March 2000–December 2016). The results showed that daily rainfalls from both 3B42V7 and 3B42RT had very weak correlations with the ground observation data; the average of the CC and the average of NSE were 0.387 and −0.152 for 3B42V7 data and 0.304 and −0.521 for 3B42RT data, respectively. The negative NSE values demonstrated that TMPA values were less accurate than the mean of observed data and were therefore very poor estimations.

Table 3. Descriptive statistics for observation rain gauge and TMPA data in daily and monthly scale.

	n	TMPA 3B42V7						TMPA 3B42RT					
		Daily Scale			Monthly Scale			Daily Scale			Monthly Scale		
		Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
CC	29	0.510	0.320	**0.387**	0.959	0.833	**0.896**	0.395	0.216	**0.304**	0.900	0.731	**0.842**
NSE	29	0.207	−0.507	**−0.152**	0.884	0.593	**0.765**	0.002	−0.968	**−0.521**	0.792	0.131	**0.480**
RMSE	29	21.7	11.4	**15.1**	111.6	36.2	**66.5**	24.5	13.7	**17.3**	143.6	76.4	**96.0**
PBIAS	29	33.2	−21.5	**3.2**	33.2	−21.5	**3.2**	38.5	−18.1	**14.8**	38.5	−18.1	**14.8**

Note: n is total number of stations. RMSE unit on a daily scale is mm/day. RMSE unit on a monthly scale is mm/month.

The statistics metric for monthly scale showed a significant improvement for both 3B42V7 and 3B42RT compared to ground data (Table 3). However, the PBIAS did not change from a daily to monthly scale. Monthly 3B42V7 and monthly 3B42RT had similar CC, with an average value of 0.896 and 0.842, respectively. However, monthly 3B42V7 data greatly outperformed monthly 3B42RT data regarding NSE, RMSE, and PBIAS. Average NSE of monthly 3B42V7 was 0.765 and no single station had a value smaller than 0.5, while average NSE of monthly 3B42RT was only 0.480. The monthly CC and NSE scores of 3B42V7 compared to ground data in this case study were very similar to the results of monthly 3B43 data compared to observations in the same basin [26]. Average RMSE of monthly 3B42V7 was 66.5 mm/month, equivalent to 30% less than the figure of monthly 3B42RT. Average PBIAS of monthly 3B42V7 was approximately 5 times less than that of monthly 3B42RT, with values of 3.2% and 14.8% respectively. The positive of PBIAS also indicated that both TMPA products overestimated compared to ground observation data. This finding was consistent with the study at the Black Volta Basin of West African countries [17] or Pahang River Basin of Malaysia [27], but it was contrary to the study in Iran [54]. It should be mentioned that although belonging to the same South East Asia region, the 3B42V7 data over the Red–Thai Binh River Basin performed better than that for Malaysia's basin as the PBIAS of 3B42V7 for Malaysia's basin was up to 26.95% on average [27].

We calculated various error metrics—CC, NSE, RMSE, and PBIAS. However, for TMPA's spatial performance purpose, we only showed the spatial PBIAS score distribution. There were two reasons for this: (1) PBIAS is recommended in water resources planning projects because the overall difference

between observed and estimated values is a criteria of paramount importance [55]; (2) PBIAS is precisely aimed at defining a poor model performance and has immense variation between seasons [56].

Looking at the PBIAS distribution, the PBIAS of 3B42V7 data mostly ranged within ±10%, while the PBIAS of 3B42RT data mostly fell in the range of 10–40% (Figure 3). The poor performance of 3B42RT data was observed at the center of the Red–Thai Binh River Basin. Moderate performances for both TMPA products were found at the northwestern mountainous area between Vietnam and Chinese border as well as the northeast coastal area.

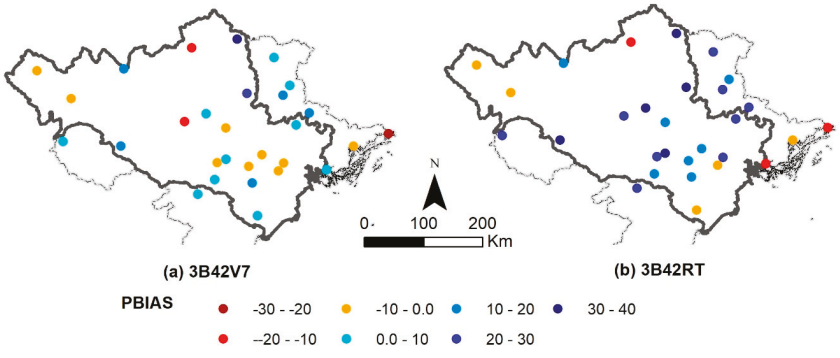


Figure 3. Percentage bias (PBIAS) score's spatial performance of TMPA products (**a**) 3B42V7 and (**b**) 3B42RT against observation data on both daily and monthly scales from March 2000 to December 2016 over Red–Thai Binh River Basin. The grey line is the Red–Thai Binh River Basin boundary within the Vietnam territory.

4.1.2. Dry and Wet Season Assessment

Table 4 presents the performances of both daily and monthly 3B42V7 and 3B42RT during the dry season (November–April) and wet season (May–October) over the Red–Thai Binh River Basin. Generally, 3B42V7 data was better than 3B42RT data in all statistical metrics compared to ground stations, especailly NSE, RMSE, and PBIAS. For example, monthly 3B42V7 had moderate NSE metric compared to ground observation, with averages of 0.586 and 0.566 in the dry season and wet season, respectively. In contrast, the figures of monthly 3B42RT were quite low, with 0.199 and 0.009, respectively. Interestingly, although RMSE of daily 3B42V7 during both dry and wet seasons were quite similar to those of daily 3B42RT, aggregation daily 3B42V7 to monthly was significantly less than monthly 3B42RT during both dry and wet seasons, with a reduction of approximately 30% for each. Regarding PBIAS, 3B42V7 and 3B42RT had almost the same bias during the dry season; however, in the wet season, 3B42V7 had significantly low PBIAS, with a value of 6.1% compared to 20.7% PBIAS of 3B42RT. In regard to the dry and wet seasons, although CC and NSE were slightly higher during the dry season than during the wet season, it was not clearly evident. On the other hand, RMSE during the dry season was observed to be much smaller than during the wet season. This can be explained as dry season receives a small amount of rainfall (10–15% of total annual rainfall), and its rainfall variation is not high as the fluctuation during the wet season.

Both TMPA products showed overall negative PBIAS values during the dry season and overall positive PBIAS values during the wet season, indicating overall underestimations during the dry season and overall overestimations during the wet season. 3B42V7 was observed to underestimate ground observation at 20 out of 29 stations, and this figure was 24 out of 29 stations for 3B42RT (Figure 4). When we used scatter plot to compare monthly dry season of TMPA products and that of ground observation (data not shown), we found that TMPA products reported zero values in many months. The wrong no-rainfall reported by TMPA data was also found at Chindwin River Basin of

Myanmar [57]. The underestimation of TMPA rainfall during the dry season was consistent with previous studies in Southwestern of China [58]. On the other hand, during the wet season, 22 out of 29 stations experienced overestimations for 3B42V7 data and 24 out of 29 stations experienced overestimation for 3B42RT data. The northwest mountain region and the northeast coastal area were the only places where both TMPA products underestimated ground observation data during two seasons. The overestimation of rainfall during the wet season agreed with a case study in Malaysia [27] but was contrary to a study involving the southwestern region of China [58].

Table 4. Descriptive statistics for daily and monthly observation rain gauge and those of TMPA data during the dry and wet seasons.

	n	TMPA 3B42V7						TMPA 3B42RT					
		Dry Season			Wet Season			Dry Season			Wet Season		
		Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
Daily													
CC	29	0.487	0.289	**0.407**	0.494	0.264	**0.344**	0.423	0.196	**0.317**	0.364	0.141	**0.255**
NSE	29	−0.048	−0.884	**−0.325**	0.185	−0.601	**−0.201**	−0.076	−1.237	**−0.612**	−0.031	−1.107	**−0.598**
RMSE	29	8.8	5.8	**7.0**	29.1	14.7	**20.0**	10.2	6.4	**7.7**	32.8	17.9	**22.9**
PBIAS	29	18.9	−38.6	**−10.4**	37.2	−19.4	**6.1**	31.7	−47.3	**−14.1**	43.2	−13.3	**20.7**
Monthly													
CC	29	0.957	0.657	**0.827**	0.924	0.588	**0.796**	0.873	0.551	**0.718**	0.817	0.358	**0.691**
NSE	29	0.881	0.191	**0.586**	0.788	0.276	**0.566**	0.672	−0.381	**0.199**	0.618	−0.845	**0.009**
RMSE	29	44.9	16.8	**29.0**	152.8	48.0	**88.7**	57.1	26.7	**41.0**	198.3	100.6	**128.6**
PBIAS	29	18.9	−38.6	**−10.4**	37.2	−19.4	**6.1**	31.7	−47.3	**−14.1**	43.2	−13.3	**20.7**

Note: n is the total number of stations. RMSE unit on a daily scale is mm/day. RMSE unit on a monthly scale is mm/month.

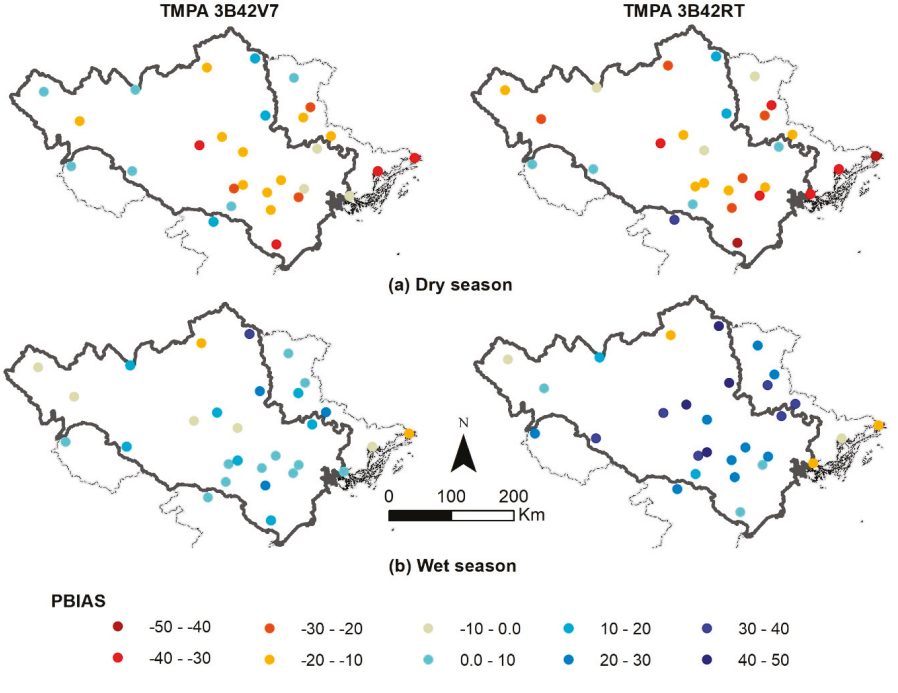


Figure 4. PBIAS score's spatial performance of TMPA rainfall data against observation data during (**a**) the dry and (**b**) the wet season from March 2000 to December 2016 over the Red–Thai Binh River Basin. The grey line is the Red–Thai Binh River Basin boundary within Vietnam territory.

Although they had a generally positive PBIAS score, TMPA products seemed to underestimate large rainfall amounts. One possible explanation for this could be their spatial resolution. With a rather low 0.25° spatial resolution (approximately 25 km), rainfall observed in a grid was averaged over about 625 km^2. However, rainfall can vary dramatically even with a few kilometers, and the resolution of TMPA products are often unable to pick up these differences. If we consider the complexity of terrain, this variation can be harder to estimate. Additionally, many convective storms can have a rapid evolution that a satellite will often not be able to observe accurately [59].

4.1.3. Rainfall Detection Assessment

The capacity of 3B42V7 and 3B42RT data regarding rainfall detection over the Red–Thai Binh River Basin from March 2000 to December 2016 is presented in Figure 5. Generally, the detection capacity of daily TMPA products during the wet season was much better than during the dry season, and 3B42V7 data had slightly better score than 3B42RT data. This may be associated with the temporal resolution of TMPA data as short-duration rainfall events are a typical characteristic of the dry season. Indeed, with 3-hourly products, it is easy for TMPA to miss events lasting less than this figure. On the other hand, TRMM is meant to capture and estimate convective precipitation rather than other types because of its on-board sensors. In Vietnam, precipitation is generally associated with heavier storms and cloud coverage during the wet season [60], meaning the precipitation is more likely to be detected. In contrast, in the dry season, there will be much lighter rainfall with less cloud coverage and convection, meaning that it will be more difficult to detect [59].

The POD for the whole daily TMPA data was stable over the years, with the average values of about 0.61 and 0.58 for 3B42V7 and 3B42RT, respectively. The POD scores for the daily time series in the wet season period were higher, with average values of 0.71 and 0.69 for 3B42V7 and 3B42RT, respectively. The POD scores of 3B42V7 and 3B42RT for the daily time series in the dry season period were typically low, with average values of 0.32 and 0.30, respectively. In the year 2012, the POD scores during the dry season were the lowest at about 0.2. The FAR of the daily time series and the daily values in wet season were moderate, with average values of 0.37 and 0.36 corresponding to 3B42V7 and 0.40 and 0.38 corresponding to 3B42RT. However, the FAR of the time series in the dry season was high, with an average of 43% of 3B42V7 rainfall prediction being wrong (FAR = 0.43). The wrong prediction of 3B42RT was even more than that of 3B42V7, with average FAR being 0.50. Interestingly, FAR scores of 3B42V7 and 3B42RT had great fluctuation over the years, reaching 0.6 and 0.62 in the year 2000, respectively, but the wrong prediction was reduced to only 0.30 for 3B42V7 and 0.42 for 3B42RT in the year 2014. The POFD scores were moderate for both TMPA products, with average values of 0.15 and 0.16 for 3B42V7 and 3B42RT, respectively. The POFD scores during the dry season were relatively low, with all years reporting values less than 0.1 for both TMPA products. During the wet season, POFD scores were higher than those of the dry season, with a range of 0.2–0.3. The CSI scores showed that there was no single year during the study time where the CSI scores of both 3B42V7 and 3B42RT were over 0.5. During the wet season, the average CSI values were around 0.52 and 0.50 for 3B42V7 and 3B42RT, respectively. Regarding the dry season, the CSI were only about 0.24 and 0.21 for each TMPA product, and the lowest CSI scores in the dry season were observed in 2006 and 2012.

As CSI combines different aspects of POD and FAR to give an overall assessment of TMPA performance, we used this metric to investigate the detection metric of TMPA products on the spatial scale (Figure 6). The lowland central part of the basin experienced the worst CSI score, while the northwestern mountainous part of the basin had moderate CSI score (>0.5). The better detection capacity at high elevation region than the lower land area was consistent with the study in Malaysia's basin [27].

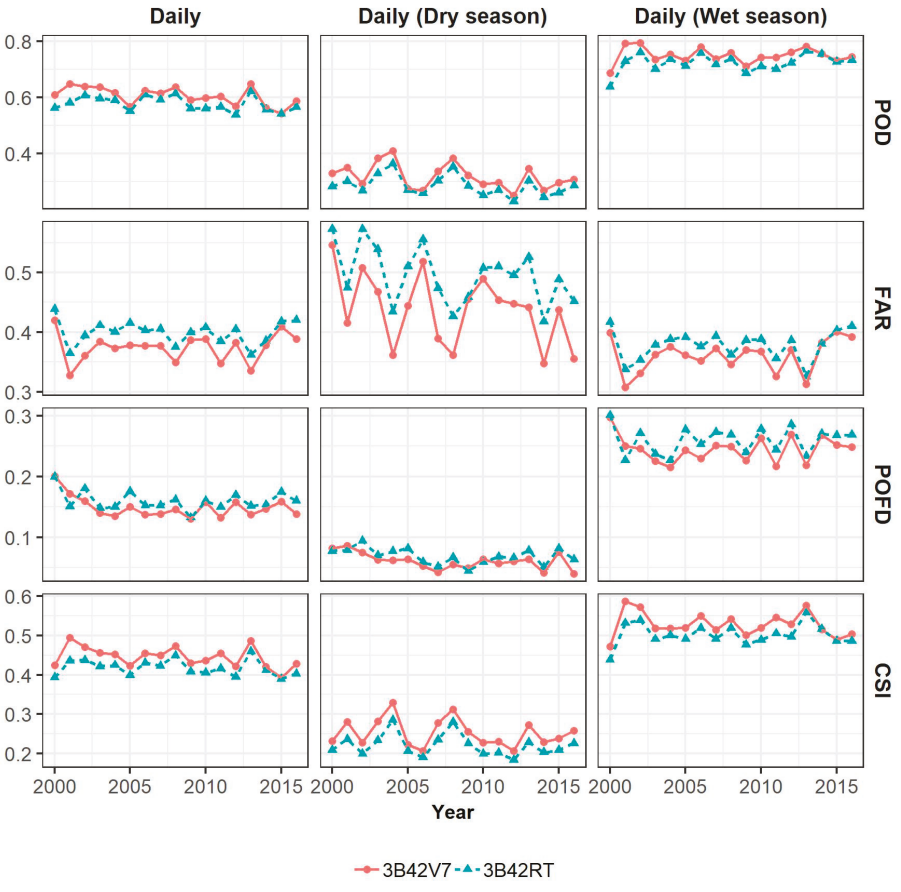


Figure 5. Average rainfall detection measurement of TMPA 3B42V7 and TMPA 3B42RT over the Red–Thai Binh River Basin from March 2000 to December 2016.

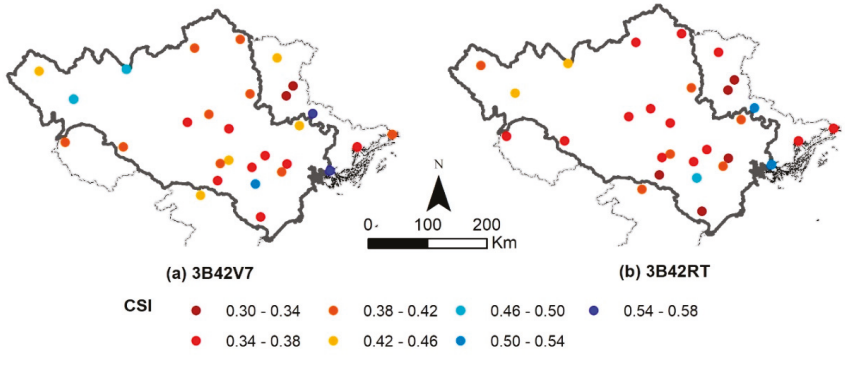


Figure 6. Critical success index (CSI) score's spatial performance of TMPA rainfall data against observation data from March 2000 to December 2016 over the Red–Thai Binh River basin. The grey line is the Red–Thai Binh River Basin boundary within Vietnam territory.

4.1.4. Rainfall Intensity Analysis

The rainfall frequency distributions of ground observations, 3B42V7, and 3B42RT over the Red–Thai Binh River Basin are presented in Figure 7. Generally, rainfall intensity of both TMPA products followed the intensity of ground observations for the whole time series. Based on ground observation data, no rainfall (≤0.6 mm/day) accounted for 68.8% of total rainfall events, and 3B42V7 and 3B42RT data had similar figures. During the dry season, low rainfall intensity (0.6–6 mm/day) detected by TMPA datasets were relatively low (4.8% and 5.3% corresponding to 3B42V7 and 3B42RT) compared to the figure from ground measurement (13.1%). However, the no rainfall (≤0.6 mm/day) detected in the dry season was a different situation. The 3B42V7 estimated 86.4% of daily rainfall events during this season as no rainfall. Similarly, 88% of rainfall events during the dry season were considered as no rainfall events by 3B42RT. In contrast, the observations data only reported a figure of 79%. During the wet season, the no rainfall events by 3B42V7 and 3B42RT were relatively low (approximately 52% for both products), while the figure for observation data was nearly 60%. Regarding high rainfall events (50–100 mm/day) and heavy rainfall events (>100 mm/day), TMPA products had a high accuracy of detecting these, with the PDFs of both TMPA products almost the same as those of observation. The slight underestimation of low rainfall event (0.6–6 mm/day) was contrary to the overestimation conclusion of this rainfall intensity in a case study in Singapore [36].

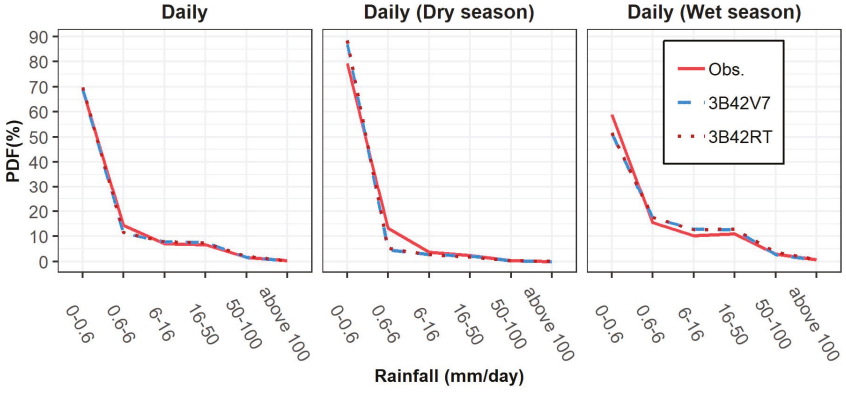


Figure 7. Average probability density function (PDF) of ground observation, TMPA 3B42V7, and TMPA 3B42RT for rainfall in daily, daily (dry season), and daily (wet season) over the Red–Thai Binh River Basin from March 2000 to December 2016.

As no rainfall and low rainfall intensity during the dry season and wet season experience significant differences between ground observations and TMPA data, we exploited the differences by analyzing seasonal spatial low rainfall and light rainfall's intensity of TMPA products. PDF differences between each TMPA data and ground observation were calculated and are presented in Figure 8. The 3B42V7 and 3B42RT data had similar characteristics, which overestimated no rainfall during the dry season (10–15%) and low rainfall intensity during the wet season (0–5%). On the other hand, the TMPA products underestimated no rainfall during the wet season (10–13%) and low rainfall intensity during the dry season (10–15%). It was noticed that the above characteristics occurred similarly for areas throughout the basin and were not specific to a typical region.

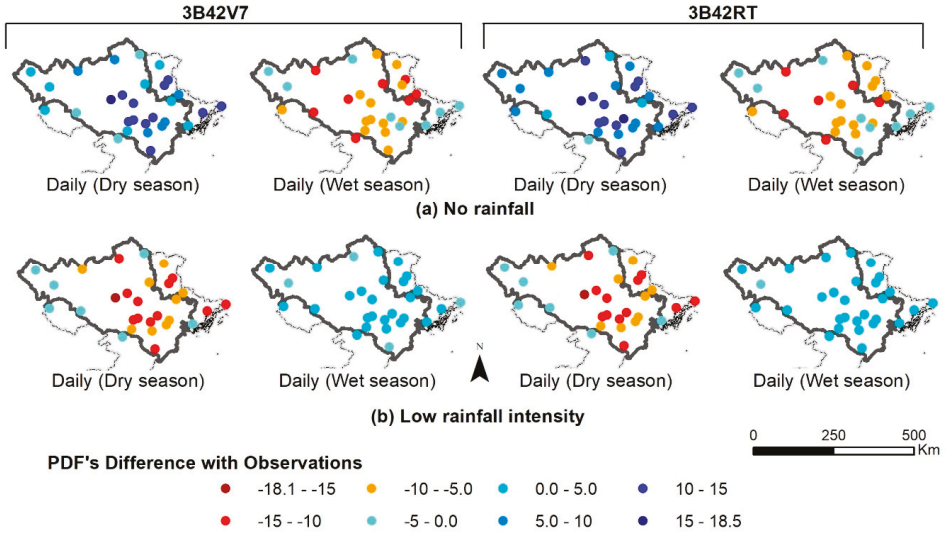


Figure 8. Percentage difference of PDF between TMPA 3B42V7, TMPA 3B42RT, and observation at (**a**) no rainfall intensity (0–0.6 mm/day) and (**b**) low rainfall intensity (0.6–6 mm/day) over the Red–Thai Binh River Basin from March 2000 to December 2016.

4.2. Development of Bias Correction Model Using Climatology–Topography Characteristics-Based Linear-Scaling (LS) Approach

4.2.1. Correlation Analysis between Climatology–Topography Characteristics and Correction Factors of LS Approach

In the LS approach, the correction factor is an important key to adjust satellite data closely to observation. Correction factors between TMPA products and observations were calculated for each month. In total, we had 12 group correction factors for 3B42V7 and 3B42RT data. Tables 5 and 6 present the relationship between correction factors in each month with climatology–topography characteristics. Based on the significant levels of the correlation coefficient, we found that topographical characteristics (LONG, LAT, and ELEV) were often associated with correction factors during dry months (except for April), while climate characteristics (AR, SDR, and NRD) were often linked with correction factors during wet months. A larger correction factor indicates larger error between satellite data and observations. ELEV (elevation) had a significant inverse relationship with the correction factor, meaning satellite data at higher elevation areas probably had a smaller error with observations compared to lower areas. This result agreed with an observation in Iran that compared 3B43V7 with rain gauge over this country [25]. Similarly, LAT (latitude) also had significant negative relationship with the correction factor. This meant that the higher the latitude area, the smaller was the satellite error. The frequency occurrence of clouds can affect the accuracy of satellite rainfall estimation [61], and NRD (a number of the rainy days) is a variable that reflects this frequency. As the number of rainy days had significant correlations with the correction factors with negative values, it seemed that the higher the number of rainy day stations, the more error of satellite-based rainfall there were. In addition, from Tables 5 and 6, AR (annual rainfall rate) and SDR (standard deviation of rainfall) had significant positive correlations with the correction factors. This means the higher the rainfall rate area, the higher was the correction factor, implying a more substantial error. This feature was the same as previous literature [62]. As a result, the correction factor for each month could be estimated from significant climatology–topography candidates.

Table 5. Correlation coefficient between correction factors of TMPA 3B42V7 against climatology–topography characteristics.

	LONG	LAT	ELEV	AR	SDR	NRD
CF_1	**−0.46** **	−0.32	−0.35	0.18	0.27	−0.21
CF_2	**0.47** **	**−0.61** **	**−0.48** **	0.09	0.29	−0.02
CF_3	**0.43** *	**−0.58** **	**−0.42** *	−0.03	0.11	0.07
CF_4	0.01	0.00	−0.11	**0.52** **	**0.47** *	**−0.57** **
CF_5	0.01	0.07	−0.02	**0.66** **	**0.56** **	**−0.65** **
CF_6	−0.04	0.02	−0.09	**0.63** **	**0.45** *	**−0.46** *
CF_7	0.03	0.12	−0.08	**0.69** **	**0.60** **	**−0.44** *
CF_8	0.06	−0.16	−0.34	**0.53** **	**0.54** **	**−0.50** **
CF_9	−0.29	0.17	0.25	**0.52** **	**0.50** **	**−0.58** **
CF_{10}	0.26	**−0.42** *	**−0.40** *	**0.58** **	**0.68** **	**−0.47** *
CF_{11}	0.29	−0.05	**−0.47** **	0.31	0.33	−0.32
CF_{12}	**0.39** *	**−0.39** *	**−0.41** *	0.12	0.23	−0.17

Note: * 0.05 significant level; ** 0.01 significant level

Table 6. Correlation coefficient between correction factors of TMPA 3B42RT against climatology–topography characteristics.

	LONG	LAT	ELEV	AR	SDR	NRD
CF_1	−0.02	**−0.50** **	−0.10	0.29	0.28	−0.30
CF_2	**0.48** **	**−0.69** **	**−0.56** **	−0.03	0.16	0.23
CF_3	0.25	**−0.55** **	**−0.42** *	−0.06	0.01	0.05
CF_4	−0.15	0.22	−0.16	**0.61** **	**0.43** *	**−0.57** **
CF_5	−0.22	0.37	−0.03	**0.69** **	**0.47** *	**−0.59** **
CF_6	−0.28	**0.44** *	0.02	**0.68** **	**0.39** *	**−0.53** **
CF_7	0.10	−0.13	−0.30	**0.61** **	**0.56** **	−0.18
CF_8	**0.48** **	−0.30	**−0.52** **	**0.58** **	**0.74** **	−0.23
CF_9	30	0.04	−0.03	**0.54** **	**0.69** **	−0.34
CF_{10}	0.19	**−0.42** *	−0.34	**0.57** **	**0.66** **	−0.35
CF_{11}	**0.55** **	−0.22	**−0.51** **	**0.46** *	**0.62** **	−0.24
CF_{12}	**0.64** **	−0.37	**−0.45** *	0.23	**0.43** *	−0.16

Note: * 0.05 significant level; ** 0.01 significant level.

4.2.2. Multiple Linear Model Development to Estimate Correction Factors

As climatology–topography characteristics have various units, before building the multiple linear regression models for correction factors, we made it dimensionless for all input climatology–topography data by scaling them to a range [0.1, 0.9]. The multiple linear models for the correction factors of 3B42V7 and 3B42RT are presented in Tables 7 and 8. All p-values were smaller than 0.5, indicating that sets of linear models using climatology–topography characteristics could well predict correction factors.

Table 7. Multiple linear models to predict correction factors of TMPA 3B42V7 data.

Formulas	CC	p-Value
$CF_1 = 1.004 * \text{LONG} + 0.947$	0.446	0.045
$CF_2 = 0.736 * \text{LONG} - 2.331 * \text{LAT} - 1.347 * \text{ELEV} + 3.256$	0.779	<0.001
$CF_3 = 0.242 * \text{LONG} - 1.298 * \text{LAT} - 0.735 * \text{ELEV} + 2.200$	0.768	<0.001
$CF_4 = 0.103 * \text{AR} + 0.134 * \text{SDR} - 0.262 * \text{NRD} + 0.921$	0.604	0.003
$CF_5 = 0.046 * \text{AR} + 0.218 * \text{SDR} - 0.404 * \text{NRD} + 1.068$	0.748	0.001
$CF_6 = 1.019 * \text{AR} - 0.528 * \text{SDR} + 0.116 * \text{NRD} + 0.662$	0.712	0.003
$CF_7 = 0.937 * \text{AR} - 0.170 * \text{SDR} + 0.140 * \text{NRD} + 0.615$	0.733	0.002
$CF_8 = -0.351 * \text{AR} + 0.544 * \text{SDR} - 0.393 * \text{NRD} + 1.093$	0.694	0.006
$CF_9 = -0.491 * \text{AR} + 0.636 * \text{SDR} - 0.563 * \text{NRD} + 1.182$	0.687	0.006
$CF_{10} = -0.286 * \text{LAT} - 0.291 * \text{ELEV} - 0.665 * \text{AR} + 0.965 * \text{SDR} - 0.678 * \text{NRD} + 1.488$	0.840	<0.001
$CF_{11} = -0.755 * \text{ELEV} + 1.373$	0.600	0.003
$CF_{12} = 0.437 * \text{LONG} - 1.422 * \text{LAT} - 1.202 * \text{ELEV} + 2.671$	0.564	0.038

Note: p-value shows significant level between predicted correction factors using multiple linear models and calculated correction factors.

Table 8. Multiple linear models to predict correction factors of TMPA 3B42RT data.

Formulas	CC	*p*-Value
$CF_1 = -1.537 * \text{LAT} + 2.582$	0.501	0.041
$CF_2 = 1.102 * \text{LONG} - 3.343 * \text{LAT} - 2.204 * \text{ELEV} + 4.038$	0.875	<0.001
$CF_3 = -2.871 * \text{LAT} - 2.045 * \text{ELEV} + 4.432$	0.748	0.001
$CF_4 = 1.041 * \text{AR} - 0.515 * \text{SDR} - 0.003 * \text{NRD} + 0.672$	0.720	0.003
$CF_5 = 1.197 * \text{AR} - 0.687 * \text{SDR} - 0.082 * \text{NRD} + 0.545$	0.761	<0.001
$CF_6 = 0.215 * \text{LAT} + 1.524 * \text{AR} - 0.909 * \text{SDR} + 0.386 * \text{NRD} + 0.147$	0.875	<0.001
$CF_7 = 0.605 * \text{AR} + 0.112 * \text{SDR} + 0.696$	0.608	0.002
$CF_8 = 0.196 * \text{LONG} - 0.241 * \text{ELEV} - 0.011 * \text{AR} + 0.580 * \text{SDR} + 0.676$	0.877	<0.001
$CF_9 = -0.311 * \text{AR} + 1.037 * \text{SDR} + 0.681$	0.712	0.003
$CF_{10} = -0.664 * \text{AR} + 0.429 * \text{SDR} + 0.587 * \text{SD} + 0.948$	0.674	0.007
$CF_{11} = 1.233 * \text{LONG} - 0.638 * \text{ELEV} + 0.720 * \text{AR} + 0.616 * \text{SDR} + 0.250$	0.838	<0.001
$CF_{12} = 2.233 * \text{LONG} - 0.978 * \text{ELEV} + 0.936 * \text{SDR} + 0.842$	0.729	0.002

Note: *p*-value shows significant level between predicted correction factors using multiple linear models and calculated correction factors.

4.2.3. Calibration and Validation of the CTLS Bias Correction Approach

Table 9 compares the TMPA products before and after using the LS and CTLS approaches against the observations on a daily scale. Both calibration and validation data showed that LS and CTLS performed very well in reduction PBIAS scores but had moderate performances regarding NSE scores, slight improvements in RMSE scores, and almost no change in CC scores. Moreover, the linear-scaling model seemed to reduce errors better for 3B42RT data compared to that for 3B42V7 data. The reason for this may be that 3B42V7 data had already passed through the correction stage before the online public, meaning other bias correction approaches did not improve this product's quality significantly. The good performances during calibration and validation stations of the CTLS approach indicated that empirical correction factors calculated by climatology and topography characteristics could be applied for the satellite-based data bias correction process throughout the Red–Thai Binh River Basin.

Table 9. The average performance of calibration and validation for climatology–topography-based linear-scaling approach (CTLS) with TMPA 34B42V7 and TMPA 3B42RT on a daily scale.

	Before Bias Correction				LS				CTLS			
	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS
3B42V7												
Calibration	0.389	−0.147	15.2	3.0	0.389	−0.130	15.2	1.1	0.389	−0.119	15.2	0.8
Validation	0.378	−0.175	14.8	4.0	0.375	−0.153	14.7	1.5	0.372	−0.165	14.8	1.7
3B42RT												
Calibration	0.303	−0.509	17.3	13.5	0.309	−0.299	16.3	−0.2	0.306	−0.342	16.6	2.2
Validation	0.307	−0.565	17.1	19.4	0.308	−0.299	15.8	0.8	0.300	−0.409	16.4	7.4

Regarding bias correction models on a monthly scale, similar results were observed as the daily scale, with a significant reduction in PBIAS scores after bias correction (Table 10). Moreover, the NSE scores of corrected monthly 3B42RT improved profoundly compared to those before bias correction. Before applying bias correction, the average monthly NSE for calibration and validation stations for 3B42RT data were 0.488 and 0.447, respectively. After using the LS approach, these figures improved to 0.734 and 0.713, respectively. Also, the empirical CTLS approach had considerable monthly NSE improvement, with values of 0.677 and 0.642 corresponding to calibration and validation stations.

Table 11 presents the performance of TMPA products regarding the PBIAS score before and after bias correction using LS and CTLS during the dry and wet seasons. The wet season seemed to benefit from bias correction more than the dry season. Using the LS approach, PBIAS scores for both 3B42V7 and 3B42RT were equal to 0, while the figures for the dry season were up to 10%. With the CTLS approach, PBIAS scores during the wet season also observed a significant decrease, ranging from 0.07% to 4.55%. During the dry season, highly positive PBIAS scores (up to 24%) were observed, indicating a high overestimation of dry season after bias correction.

Table 10. The average performance of calibration and validation for CTLS with TMPA 34B42V7 and TMPA 3B42RT on a monthly scale.

	Before Bias Correction				LS				CTLS			
	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS
3B42V7												
Calibration	0.899	0.767	66.7	3.0	0.904	0.816	59.5	1.1	0.899	0.799	62.1	0.8
Validation	0.881	0.755	65.8	4.0	0.883	0.778	63.2	1.5	0.873	0.755	65.9	1.7
3B42RT												
Calibration	0.843	0.488	96.1	13.5	0.866	0.734	71.8	−0.2	0.850	0.677	78.8	2.2
Validation	0.838	0.447	95.9	19.4	0.854	0.713	71.3	0.8	0.831	0.642	79.3	7.4

Table 11. The average PBIAS score's performance of calibration and validation for CTLS with TMPA 34B42V7 and TMPA 3B42RT during the dry and wet seasons.

	Before Bias Correction		LS		CTLS	
	Dry Season	Wet Season	Dry Season	Wet Season	Dry Season	Wet Season
3B42V7						
Calibration	−10.32	5.75	6.97	0.00	3.37	0.51
Validation	−10.92	19.29	9.63	0.00	12.67	0.07
3B42RT						
Calibration	−14.48	7.32	−0.58	0.00	5.59	1.71
Validation	−12.75	26.23	5.59	0.00	24.32	4.55

Table 12 shows the CSI scores for TMPA products against observations before and after bias correction using LS and CTLS during daily, daily (dry season), and daily (wet season). Generally, there was no significant change in CSI scores after bias correction compared to before bias correction for both products and for both bias correction approaches. By analyzing the intensity metric before and after doing bias correction (data not shown), we also obtained the same results as that of detection metric, i.e., after doing bias correction, there was no significant change in the intense rainfall for TMPA products.

Table 12. Average CSI score's performance of calibration and validation for CTLS with TMPA 34B42V7 and TMPA 3B42RT for daily, daily (dry season), and daily (wet season).

	Before Bias Correction			LS			CTLS		
	Daily	Daily (Dry Season)	Daily (Wet Season)	Daily	Daily (Dry Season)	Daily (Wet Season)	Daily	Daily (Dry Season)	Daily (Wet Season)
3B42V7									
Calibration	0.450	0.258	0.531	0.449	0.259	0.529	0.448	0.259	0.529
Validation	0.429	0.225	0.519	0.428	0.225	0.516	0.428	0.226	0.518
3B42RT									
Calibration	0.422	0.226	0.505	0.422	0.226	0.505	0.422	0.226	0.505
Validation	0.402	0.202	0.492	0.402	0.205	0.492	0.402	0.206	0.492

5. Conclusions

TMPA products are recommended for wide use over the tropical and subtropical regions due to their high temporal–spatial resolution. Therefore, this study carried out a comparison and bias correction between TMPA 3B42V7 and TMPA 3B42RT data and 29 ground observations over the lower part of the Red–Thai Binh River Basin from March 2000 to December 2016.

Based on various error metrics—CC, NSE, RMSE, and PBIAS—we compared 3B42V7 and 3B42RT against observations at different scales, including daily and monthly scales and dry and wet seasons. Our analysis showed that both products had relatively weak relationships with observations on a daily scale but this significantly improved on a monthly scale. Except for the CC score, 3B42V7 data was considerably better than 3B42RT data in the rest of the error metrics on a monthly scale. In addition, 3B42V7 data showed better performance than 3B42RT data during both dry and wet seasons, especially

regarding NSE and PBIAS measurements. Both products showed overall underestimations during the dry season and overestimations during the wet season. Spatial analysis of the PBIAS score indicated significant bias of TMPA products at the lowland area of the Red–Thai Binh River Basin, while the northwestern mountainous area and the northeast coastal area had low PBIAS for both products.

The comparison between 3B42V7 and 3B42RT was also viewed from a different angle using detection metrics—POD, FAR, POFD, and CSI—against observations on daily time series, daily time series in the dry season, and daily time series in the wet season. In this case, the 3B42V7 showed a slightly better performance compared to 3B42RT for the metrics mentioned. Both products had better detection metrics in the wet season compared to the dry season. Spatial CSI score distribution showed that the lowland area of the central basin had the lowest score compared to other parts.

From the perspective of the assessment on rainfall intensity on daily time series for the dry and wet seasons, it was found out that 3B42RT performed the same as 3B42V7 data. Both products overestimated no rainfall ($\leq$0.6 mm/day) during the dry season and underestimated rainfall intensity during the wet season. The overestimation and underestimation compromised the daily time series for the dry and wet seasons, meaning the frequency distributions of no rainfall events were almost the same for TMPA products and ground observations. On the other hand, TMPA products underestimated low rainfall intensity (0.6–6 mm/day) during the dry season and overestimated rainfall intensity during the wet season. The underestimation of low rainfall was more significant than the overestimation, resulting in a slightly lower rainfall estimation by TMPA products at the daily time series compared to observations.

In addition, we used the LS approach to do bias correction for 3B42V7 and 3B42RT products. In this approach, the correction factor is an important key to adjust satellite rainfall data closely to observations. We found that the correction factors of the LS approach were associated with climatology–topography characteristics. Therefore, a set of multiple linear regression models was developed to predict correction factors from climatology–topography characteristics for 3B42V7 and 3B42RT. After bias correction using LS and CTLS approaches, corrected TMPA products showed significant improvement compared to the results before bias correction, especially for the 3B42RT dataset with PBIAS and NSE scores. However, we found that both bias correction approaches did not improve the TMPA products significantly on other measurement scores.

In conclusion, 3B42V7 and 3B42RT data should be a good alternative source for a wide range of hydrological purposes on a monthly scale. The 3B42V7 data is also a good source for typical analysis of dry and wet seasons, although these datasets should be used with caution for daily scale purposes. The post-TMPA products after using climatology–topography characteristics are promising sources, especially for total water resource estimation.

The biggest advantage of the LS approach was to reduce PBIAS score; however, other error scores remained almost the same. Future studies may merge satellite-based and ground-based rainfall product to further improve rainfall product quality [63]. The finding of this paper gives an overview of the capacity of TMPA products in the lower part of the Red–Thai Binh River Basin regarding water resource application and provides a simple bias correction that can be used to improve the correctness of TMPA products. Additionally, the study is beneficial for regions, such as Vietnam, that are seeking alternative rainfall sources. The reason for this is that approximately 60% of Vietnam's water resources come from abroad, and hydro-climatology acquisition from upstream countries faces many challenges due to limited administration interaction [64].

Author Contributions: Conceptualization, H.M.L., and V.L.; Methodology, H.M.L.; Validation, J.R.P.S., D.D.B.; Resources, H.M.L., and J.R.P.S.; Writing—Original Draft Preparation, H.M.L.; Writing—Review & Editing, H.M.L. and J.R.P.S.; Supervision, V.K. and D.D.B.; Project Administration, V.L. and J.D.B.; Funding Acquisition, V.L. and J.D.B.

Acknowledgments: The authors wish to thank the NASA SERVIR project (Program Manager Nancy Searby) for the funding. The authors thank the hydrological modeling group of the National Central for Water Resources Planning and Investigation (NAWAPI) for providing essential support for data collection. The authors also thank anonymous reviewers for their valuable comments to improve the manuscript's quality.

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

References

1. Kumar, B.; Lakshmi, V. Accessing the capability of TRMM 3B42 V7 to simulate streamflow during extreme rain events: Case study for a Himalayan River Basin. *J. Earth Syst. Sci.* **2018**, *127*, 27. [CrossRef]
2. Brutsaert, W. *Hydrology: An Introduction*; Cambridge University Press: Cambridge, UK, 2005.
3. 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]
4. Kidd, C. Satellite rainfall climatology: A review. *Int. J. Climatol.* **2001**, *21*, 1041–1066. [CrossRef]
5. Rana, S.; McGregor, J.; Renwick, J. Precipitation seasonality over the Indian subcontinent: An evaluation of gauge, reanalyses, and satellite retrievals. *J. Hydrometeorol.* **2015**, *16*, 631–651. [CrossRef]
6. Xie, P.; Arkin, P.A. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Clim.* **1996**, *9*, 840–858. [CrossRef]
7. Plengsaeng, B.; Wehn, U.; van der Zaag, P. Data-sharing bottlenecks in transboundary integrated water resources management: A case study of the Mekong River Commission's procedures for data sharing in the Thai context. *Water Int.* **2014**, *39*, 933–951. [CrossRef]
8. Gerlak, A.K.; Lautze, J.; Giordano, M. Water resources data and information exchange in transboundary water treaties. *Int. Environ. Agreem. Polit. Law Econ.* **2011**, *2*, 179–199. [CrossRef]
9. Viglione, A.; Borga, M.; Balabanis, P.; Blöschl, G. Barriers to the exchange of hydrometeorological data in Europe: Results from a survey and implications for data policy. *J. Hydrol.* **2010**, *394*, 63–77. [CrossRef]
10. García, L.; Rodríguez, D.; Wijnen, M.; Pakulski, I. *Earth Observation for Water Resources Management: Current Use and Future Opportunities for the Water Sector*; World Bank Publications: Washington, DC, USA, 2016.
11. 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**. [CrossRef]
12. 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]
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. Funk, C.C.; Peterson, P.J.; Landsfeld, M.F.; Pedreros, D.H.; Verdin, J.P.; Rowland, J.D.; Romero, B.E.; Husak, G.J.; Michaelsen, J.C.; Verdin, A.P. *A Quasi-Global Precipitation Time Series for Drought Monitoring*; US Geological Survey Data Series 832; U.S. Geological Survey: Sioux Falls, SD, USA, 2014.
15. 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]
16. Ha, L.T.; Bastiaanssen, W.G.; Griensven, A.V.; Van Dijk, A.I.; Senay, G.B. Calibration of Spatially Distributed Hydrological Processes and Model Parameters in SWAT Using Remote Sensing Data and an Auto-Calibration Procedure: A Case Study in a Vietnamese River Basin. *Water* **2018**, *10*, 212. [CrossRef]
17. Adjei, K.A.; Ren, L.; Appiah-Adjei, E.K.; Odai, S.N. Application of satellite-derived rainfall for hydrological modelling in the data-scarce Black Volta trans-boundary basin. *Hydrol. Res.* **2015**, *46*, 777–791. [CrossRef]
18. Xue, X.; Hong, Y.; Limaye, A.S.; Gourley, J.J.; Huffman, G.J.; Khan, S.I.; Dorji, C.; Chen, S. Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J. Hydrol.* **2013**, *499*, 91–99. [CrossRef]
19. Sahoo, A.K.; Sheffield, J.; Pan, M.; Wood, E.F. Evaluation of the tropical rainfall measuring mission multi-satellite precipitation analysis (TMPA) for assessment of large-scale meteorological drought. *Remote Sens. Environ.* **2015**, *159*, 181–193. [CrossRef]
20. Zhang, A.; Jia, G. Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Remote Sens. Environ.* **2013**, *134*, 12–23. [CrossRef]

21. Arvor, D.; Dubreuil, V.; Ronchail, J.; Simões, M.; Funatsu, B.M. Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil). *Int. J. Climatol.* **2014**, *34*, 2622–2633. [CrossRef]
22. Cashion, J.; Lakshmi, V.; Bosch, D.; Jackson, T.J. Microwave remote sensing of soil moisture: Evaluation of the TRMM microwave imager (TMI) satellite for the Little River Watershed Tifton, Georgia. *J. Hydrol.* **2005**, *307*, 242–253. [CrossRef]
23. Li, Z.; Yang, D.; Gao, B.; Jiao, Y.; Hong, Y.; Xu, T. Multiscale hydrologic applications of the latest satellite precipitation products in the Yangtze River Basin using a distributed hydrologic model. *J. Hydrometeorol.* **2015**, *16*, 407–426. [CrossRef]
24. Tong, K.; Su, F.; Yang, D.; Hao, Z. Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *J. Hydrol.* **2014**, *519*, 423–437. [CrossRef]
25. Moazami, S.; Golian, S.; Kavianpour, M.R.; Hong, Y. Comparison of PERSIANN and V7 TRMM Multi-satellite Precipitation Analysis (TMPA) products with rain gauge data over Iran. *Int. J. Remote Sens.* **2013**, *34*, 8156–8171. [CrossRef]
26. Simons, G.; Bastiaanssen, W.; Ngô, L.A.; Hain, C.R.; Anderson, M.; Senay, G. Integrating global satellite-derived data products as a pre-analysis for hydrological modelling studies: A case study for the Red River Basin. *Remote Sens.* **2016**, *8*, 279. [CrossRef]
27. Zad, M.; Najja, S.; Zulkafli, Z.; Muharram, F.M. Satellite Rainfall (TRMM 3B42-V7) Performance Assessment and Adjustment over Pahang River Basin, Malaysia. *Remote Sens.* **2018**, *10*, 388.
28. Kneis, D.; Chatterjee, C.; Singh, R. Evaluation of TRMM rainfall estimates over a large Indian river basin (Mahanadi). *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2493–2502. [CrossRef]
29. Curtarelli, M.P.; Rennó, C.D.; Alcântara, E.H. Evaluation of the Tropical Rainfall Measuring Mission 3B43 product over an inland area in Brazil and the effects of satellite boost on rainfall estimates. *J. Appl. Remote Sens.* **2014**, *8*, 1–14. [CrossRef]
30. Cao, Y.; Zhang, W.; Wang, W. Evaluation of TRMM 3B43 data over the Yangtze River Delta of China. *Sci. Rep.* **2018**, *8*, 5290. [CrossRef] [PubMed]
31. Huffman, G.J. The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG); NASA/GSFC Code. Available online: https://pmm.nasa.gov/sites/default/files/document_files/TMPA-to-IMERG_transition_161025.pdf (accessed on 8 June 2018).
32. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation*; NASA/GSFC Code; NASA's Goddard Space Flight Center: Greenbelt, MD, USA, 2015; Volume 612, p. 47.
33. 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]
34. 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]
35. 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]
36. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]
37. Khan, S.I.; Hong, Y.; Gourley, J.J.; Khattak, M.U.K.; Yong, B.; Vergara, H.J. Evaluation of three high-resolution satellite precipitation estimates: Potential for monsoon monitoring over Pakistan. *Adv. Space Res.* **2014**, *54*, 670–684. [CrossRef]
38. 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]
39. WMO. *Guide to Hydrological Practices: Data Acquisition and Processing, Analysis, Forecasting and Other Applications*; World Meteorological Organization: Geneva, Switzerland, 1994.
40. Duc, D.D. Assessment current situation and possibility of exploting satellite rainfall for flood forecasting- an application for Chu River Basin. *J. Clim. Chang. Sci.* **2017**, *2*, 98–104. (In Vietnamese)

41. 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]
42. Nguyen, T.H.; Masih, I.; Mohamed, Y.A.; van der Zaag, P. Validating Rainfall-Runoff Modelling Using Satellite-Based and Reanalysis Precipitation Products in the Sre Pok Catchment, the Mekong River Basin. *Geosciences* **2018**, *8*, 164. [CrossRef]
43. NAWAPI. *Red-Thai Binh River Basin Water Resources Planning, Term of Reference*; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2017. (In Vietnamese)
44. NAWAPI. *Red-Thai Binh River Basin Water Resources Planning, Main Report*; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2017. (In Vietnamese)
45. MONRE. *Circular on National Technical Standard of Meteorological Monitoring*; 25/2012/TT-BTNMT; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2012.
46. Huffman, G.J.; Bolvin, D.T. *TRMM and Other Data Precipitation Data Set Documentation*; NASA: Greenbelt, MD, USA, 2013; Volume 28.
47. Legates, D.R.; McCabe, G.J. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]
48. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900. [CrossRef]
49. Schaefer, J.T. The critical success index as an indicator of warning skill. *Weather Forecast.* **1990**, *5*, 570–575. [CrossRef]
50. NCHMF. *Rainfall Classification of Vietnam*; NCHMF: Hanoi, Vietnam, 2000.
51. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences (International Geophysics Series; V. 91)*; Academic Press: Cambridge, MA, USA, 2006.
52. Lenderink, G.; Buishand, A.; Deursen, W.V. Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1145–1159. [CrossRef]
53. Teutschbein, C.; Seibert, J. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.* **2012**, *456*, 12–29. [CrossRef]
54. Alijanian, M.; Rakhshandehroo, G.R.; Mishra, A.K.; Dehghani, M. Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. *Int. J. Climatol.* **2017**, *37*, 4896–4914. [CrossRef]
55. NAWAPI. *Bang Giang—Ky Cung Water Resources Planning Project, Water Resources Assessment Report*; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2018. (In Vietnamese)
56. Gupta, H.V.; Sorooshian, S.; Yapo, P.O. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrol. Eng.* **1999**, *4*, 135–143. [CrossRef]
57. Yuan, F.; Zhang, L.; Win, K.W.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Assessment of GPM and TRMM Multi-Satellite Precipitation Products in Streamflow Simulations in a Data-Sparse Mountainous Watershed in Myanmar. *Remote Sens.* **2017**, *9*, 302. [CrossRef]
58. Hu, Q.; Yang, D.; Li, Z.; Mishra, A.K.; Wang, Y.; Yang, H. Multi-scale evaluation of six high-resolution satellite monthly rainfall estimates over a humid region in China with dense rain gauges. *Int. J. Remote Sens.* **2014**, *35*, 1272–1294. [CrossRef]
59. 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]
60. Nguyen, D.Q.; Renwick, J.; McGregor, J. Variations of surface temperature and rainfall in Vietnam from 1971 to 2010. *Int. J. Climatol.* **2014**, *34*, 249–264. [CrossRef]
61. Ochoa-Sánchez, A.; Pineda Ordonez, L.E.; Crespo, P.; Willems, P. Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific-Andean region of Ecuador and Peru. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 3179–3193. [CrossRef]
62. Almazroui, M. Calibration of TRMM rainfall climatology over Saudi Arabia during 1998–2009. *Atmos. Res.* **2011**, *99*, 400–414. [CrossRef]

63. Nerini, D.; Zulkafli, Z.; Wang, L.-P.; Onof, C.; Buytaert, W.; Lavado-Casimiro, W.; Guyot, J.-L. A comparative analysis of TRMM–rain gauge data merging techniques at the daily time scale for distributed rainfall–runoff modeling applications. *J. Hydrometeorol.* **2015**, *16*, 2153–2168. [CrossRef]
64. Xuan, T.T.; Tuyen, H.M.; Thai, T.T.; Dung, N.K. *Water Resources on Vietnam's River System*; Science and Technology Publisher: Hanoi, Vietnam, 2012. (In Vietnamese)

remote sensing

MDPI

Article

Evaluation and Hydrological Utility of the Latest GPM IMERG V5 and GSMaP V7 Precipitation Products over the Tibetan Plateau

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

Received: 15 November 2018; Accepted: 9 December 2018; Published: 12 December 2018

Abstract: Satellite precipitation products provide alternative precipitation data in mountain areas. This study aimed to assess the performance of the latest Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG) version 5 (IMERG V5) and Global Satellite Mapping of Precipitation version 7 (GSMaP V7) products and their hydrological utilities over the Tibetan Plateau (TP). Here, two IMERG Final Run products (uncalibrated IMERG (IMERG-UC) and gauge-calibrated IMERG (IMEEG-C)) and two GSMaP products (GSMaP Moving Vector with Kalman Filter (GSMaP-MVK) and gauge-adjusted GSMaP (GSMaP-Gauge)) were evaluated from April 2014 to March 2017. Results show that all four satellite precipitation products could generally capture the spatial patterns of precipitation over the TP. The two gauge-adjusted products were more consistent with the ground measurements than the satellite-only products in terms of statistical assessment. For hydrological simulation, IMERG-UC and GSMaP-MVK showed unsatisfactory performance for hydrological utility, while GSMaP-Gauge demonstrated comparable performance with gauge reference data, suggesting that GSMaP-Gauge can be selected for hydrological application in the TP. Our study also indicates that accurately measuring light rainfall and winter snow is still a challenging task for the current satellite precipitation retrievals.

Keywords: satellite precipitation; Tibetan Plateau; GPM; IMERG; GSMaP

1. Introduction

Precipitation is a key variable of the global water cycle and the driving force for land surface hydrological processes [1–3]. Reliable precipitation input is crucial for hydrological modeling and prediction [4]. Traditionally, rainfall observations are operated through rain gauges and ground-based weather radars. However, measurements from these in situ observational networks are restricted for the remote areas and mountain regions [5]. For instance, the Tibetan Plateau (TP) with an average elevation of over 4000 m, is the highest plateau in the world [6]. To date, precipitation observation networks are sparse or nonexistent in many parts of the TP due to the complex terrain, harsh climate, and high cost [7]. Satellite precipitation products are practical means of detecting rainfall distribution over vast regions that are poorly gauged, and provide a potential complementary source beyond ground in situ data [8,9].

Satellite precipitation products combine the advantage of the frequent sampling of infrared (IR) and more accurate precipitation estimates of passive microwave (PMW) observations and can provide quasi-global precipitation maps [10]. Since the Tropical Rainfall Measuring Mission (TRMM) started in 1997, various satellite precipitation products were made available to the public with different temporal and spatial resolutions [11–15]. This greatly facilitates hydrological prediction and water resource

management [16–18]. However, errors and uncertainties exist in satellite precipitation products due to the indirectness of distant radiance measurements to surface precipitation, especially in high latitudes and altitudes [3,19,20]. For example, Tian and Peters-Lidard [21] assessed six different TRMM-era datasets over the globe, and found high relative uncertainties over complex terrain areas, including the TP, the Rockies, and the Andes. Yong et al. [22] argued that the highly accurate IR and PMW observations still have difficulties on the TP with snow cover and complex terrain.

Recently, as a successor to the TRMM, the Global Precipitation Measurement (GPM) Core Observatory (GPM CO), carrying advanced precipitation sensors with the Ka/Ku-band (35.5 and 13.6 GHz, respectively) dual-frequency precipitation radar (DPR) and the multifrequency (10–183 GHz) GPM microwave imager (GMI), was launched on 28 February 2014 [23]. The GPM CO and another 10 or so satellite partners together provide the best possible sampling and coverage of precipitation observations [24,25], and can improve the capabilities of measuring light rain and snowfall. With such a constellation of varying precipitation sensors, the Integrated Multi-satellite Retrievals for GPM (IMERG) algorithm was developed to provide a new generation of precipitation estimates around the globe with high temporal (half-hour) and spatial ($0.1^\circ \times 0.1^\circ$) resolutions [23,26,27]. Based on the difference of data latency and accuracy, IMERG offers three distinct products: Early, Late, and Final Runs. These IMERG products take advantage of several existing precipitation retrieval algorithms, including TRMM Multi-satellite Precipitation Analysis (TMPA), Climate Prediction Center Morphing with Kalman filter (CMORPH-KF), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks and a Cloud Classification System (PERSIANN-CCS), and are expected to have less error and better performance than TRMM-era satellite precipitation products.

As the GPM era is now upon us, several recent studies evaluated the performance of IMERG data by comparing them against observations from gauges, radars, and other prior satellite precipitation products [10,28–32]. Huffman et al. [33] did a preliminary comparison between the IMERG and TMPA 3B42 over coterminous United States (US) in the months April through July 2014. They reported that IMERG has good agreement with 3B42 data (correlation of 0.88). Furthermore, Tang et al. [29] demonstrated that IMERG has better performance than 3B42V7 in almost every sub-region of mainland China, especially in relatively dry climate regions. Gebregiorgis et al. [32] investigated the degree of improvement of the IMERG by comparing the uncalibrated version 3 Late Run IMERG and version 7 TMPA-RT products. They concluded that IMERG shows substantial improvement with miss-rain and false-rain bias reduction and hit rate. Similar assessments of the IMERG were conducted in different regions and climatic conditions (e.g., Africa [34], the Central Amazon [35], and Iran [36], among many others). In summary, the earlier studies found that the IMERG is a good successor of TMPA, and is promising for hydrological and meteorological applications. Meanwhile, authors also highlighted that the continuous researching of IMERG for complex topography is needed as more IMERG data become available in the future [37–39].

The algorithm for IMERG underwent several upgrades since it was released with initial version 3. Recently, IMERG version 5 products were released in November 2017 and contain data from March 2014 to present. According to Huffman et al. [40], version 5 supersedes all prior IMERG versions, and users are recommended to shift to the new datasets as soon as practical. On the other hand, the Japanese counterpart of IMERG, the Global Satellite Mapping of Precipitation (GSMaP), operated by the Japan Aerospace Exploration Agency (JAXA), distributes global precipitation maps with different products: the near-real-time GSMaP_NRT, the standard GSMaP_MVK, the gauge-adjusted standard GSMaP_Gauge, and others [41]. All GSMaP products are available at $0.1^\circ \times 0.1^\circ$ resolution every hour and can be found at http://sharaku.eorc.jaxa.jp/GSMaP/. The latest GSMaP algorithm is version 7, which was updated on January 2017, and corresponding products are available after March 2014. Thus, it is necessary to evaluate the latest versions of IMERG and GSMaP products and offer some insight for product developers and users.

As the "Asian Water Tower" and "Third Pole" [6,42], the TP plays an important role in water supply of Asia's main river basins and global climate change. Knowing the spatial and temporal

patterns of precipitation is vital for meteorology research and water management of the TP. With the improvement in measuring light rain and falling snow, the GPM-era precipitation products seem to have vast application prospects in the TP area. Therefore, there is an urgent need to validate the newly available version of GPM-era satellite precipitation products over the TP. In this study, the performances of IMERG version 5 (IMERG V5) and GSMaP version 7 (GSMaP V7) products were evaluated in the TP through extensive statistical and hydrological validation. We expect that the results could provide useful information for algorithm developers of the newest precipitation versions, which are scheduled to be reprocessed to TRMM-era datasets in the near future. The remaining parts of this paper are organized as follows: Sections 2 and 3 describe the study area, the satellite and gauge data, and the evaluation method. The results are presented and discussed in Section 4. Finally, Section 5 summarizes major conclusions of this study.

2. Study Area and Data

2.1. Study Area

The TP, which covers an area of about 2.5×10^6 km^2, is located in southwestern China (Figure 1). Our study focuses on the area within China (73°–105°E, 25°–40°N), which encompasses six provinces, namely the entire Tibet Autonomous Region and the Qinghai Province, and parts of the Xin Jiang Uygur Autonomous Region, Gansu, Sichuan, and Yunnan Provinces. In the TP, topography is extremely complicated, but it has an overall descending tendency from west to east. Specially, the highest mountain in the world, Mount Everest, stands on the southern fringe of the TP. Due to its unique geographical location and high elevation, the TP has great influence on regional and even global energy and water cycles through thermal and dynamical forcings [43]. Climate on the TP is influenced by multiple climatic systems, which are characterized by the Indian monsoon and East Asian monsoon in the summer and Westerlies in the winter. The monsoon-dominated rainfall varies greatly in space with an obvious southeast to northwest gradient, and annual precipitation ranges from over 1500 mm to less than 100 mm [44]. In the summer, the southeast monsoon produces heavy precipitation, specifically in southeastern Tibet. In the winter, the snow brought by westerly winds is the main precipitation form in western regions [45]. The TP has an overall low annual mean temperature below 0 °C, with the warmest month in July. Overall, the annual mean temperature decreases from east to west across the TP. Correspondingly, vegetation types also exhibit a spatial transition from forests in the southeast, to temperate shrubland or meadow in the middle, to desert land in the northwest regions [46]. Many of Asia's major rivers, such as the Indus, Yangtze, Yellow, Brahmaputra, Mekong, and Salween rivers, all originate from the TP. In this study, the headwater region of the Yellow River, with the outlet of the Tangnaihai hydrological station, was selected for hydrological evaluation of the satellite precipitation products.

2.2. Satellite Data

IMERG is the level 3 multi-satellite precipitation algorithm of GPM, which combined all available constellation microwave precipitation estimates, infrared (IR) satellite estimates, and monthly precipitation gauge data, with the intention of creating a new generation of global precipitation products [23]. The IMERG system is run several times, first to produce an "early" multi-satellite product (~4 h after observation time) for giving a preliminary estimate, and successively providing "late" multi-satellite products (~12 h after observation time) as more data arrive. When the monthly gauge data is received, the "final" satellite-gauge product (~2.5 months after the observation month) is created as research [47]. For these three IMERG products, in addition to the difference in sensor data latency, the early and late estimates use climatological gauge data, while the final product is adjusted by monthly Global Precipitation Climatology Centre (GPCC) gauge data [48]. There are two precipitation data field variables embedded in each IMERG product: precipitationCal and precipitationUncal. PrecipitationCal means a multi-satellite precipitation estimate with gauge

calibration, and precipitationUncal presents the original multi-satellite precipitation estimate. For a more detailed description of IMERG algorithm, readers can refer to Huffman et al. [26,40,47]. In this study, we used both calibrated precipitation estimate and uncalibrated precipitation estimates from Final Run IMERG version 5 (hereafter referred to as "IMERG-C" and "IMERG-UC", respectively).

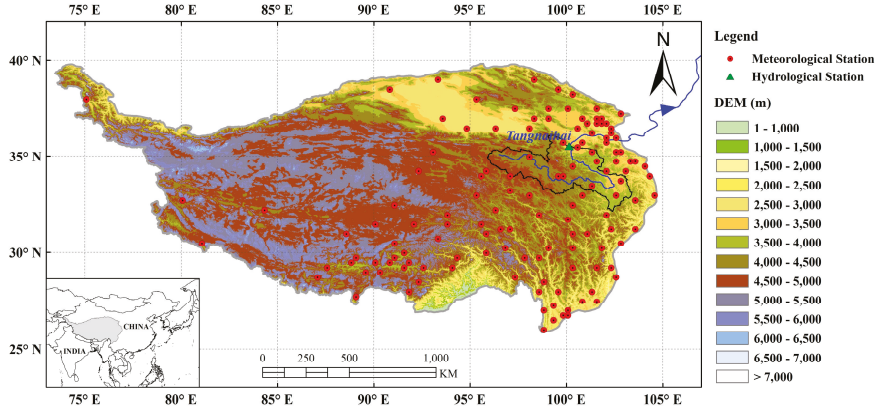


Figure 1. Map and topography of the Tibetan Plateau (TP), meteorological stations, and Tangnaihai hydrological station.

GSMaP is a satellite-based precipitation map algorithm that combines various available PMW and IR sensors aiming to develop high-precision precipitation products. The GSMaP products are produced in several steps. Firstly, the instantaneous precipitation rate is retrieved based on the PMW radiometers from different satellite platforms, including GMI, advanced microwave scanning radiometer 2 (AMSR2), TRMM Microwave Imager (TMI), special sensor microwave imager/sounder (SSMIS), advanced microwave sounding unit-A (AMSU-A), and microwave humidity sounder (MHS) [49]. Then, the gaps between PMW-based estimates are propagated using the cloud motion vectors computed from geo-IR images, and a Kalman filter approach is applied to refine the precipitation rate [50,51]. Next, the forward and backward propagated precipitation estimates are weighted and combined to generate the GSMaP-MVK product. Unlike GSMaP-MVK, GSMaP-NRT is a simplified algorithm that only considers temporary forward cloud movement to keep operability and data latency in near real time. The GSMaP-Gauge is a gauge-calibrated product that adjusts the GSMaP-MVK estimate with National Oceanic and Atmospheric Administration (NOAA)/Climate Prediction Center (CPC) gauge-based analysis of global daily precipitation [52]. In the current study, GSMaP-MVK and GSMaP-Gauge with the latest version 7 were used.

Overall, four satellite-based precipitation datasets (IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge), from April 2014 to March 2017, were selected for analysis in this study. Basic information of the four datasets is provided in Table 1. IMERG-UC and GSMaP-MVK belong to the "satellite-only" class, while IMERG-C and GSMaP-Gauge are "gauge-adjusted" products. In order to keep consistent with the ground reference dataset, both IMERG and GSMaP products were aggregated into daily readings, with $0.25^\circ \times 0.25^\circ$ resolution. Generally, satellite precipitation products perform better when scaled up to a larger spatial average and longer time periods [48,53]. Readers should keep in mind that the resampling might introduce additional uncertainties.

Table 1. Coverage and spatiotemporal resolutions of satellite precipitation products used in this study.

Product	Temporal Resolution	Spatial Resolution	Start Time	Latency	Coverage	Corrected by Gauges
IMERG-UC	0.5 h	0.1°	12 March 2014	3 months	90°N–90°S	No
IMERG-C	0.5 h	0.1°	12 March 2014	3 months	90°N–90°S	Yes (GPCC monthly)
GSMaP-MVK	1 h	0.1°	1 March 2014	3 days	60°N–60°S	No
GSMaP-Gauge	1 h	0.1°	1 March 2014	3 days	60°N–60°S	Yes (CPC daily)

2.3. Ground Gauge Data

The China Gauge-based Daily Precipitation Analysis (CGDPA), operated by the National Meteorological Information Center (NMIC) of the China Meteorological Administration (CMA), was used as the reference dataset in this study. All the gauge data used in CGDPA are manually recorded by bucket rain gauges and underwent a strict quality control procedure [54]. Using the optimal interpolation (OI) method, daily precipitation data from about 2400 gauge stations in mainland China were interpolated to obtain the gridded CGDPA product (0.25° × 0.25°). The CGDPA is considered as a high-quality ground precipitation dataset and was previously adopted to evaluate satellite precipitation products over China [55–57].

For the hydrological evaluation, the daily meteorological observations from 2009 to 2015 at the upper Yellow River basin were obtained from CMA, including daily maximum and minimum temperature, and daily average wind speed. The daily streamflow data at the Tangnaihai hydrological station from 2009 to 2015 were primarily collected from the Chinese Hydrology Almanac, which is published by the Hydrological Bureau of the Ministry of Water Resources in the People's Republic of China.

2.4. Geographical Data

Geographical data, such as underlying surface information required by the Variable Infiltration Capacity (VIC) model, including soil texture, topography, and vegetation types, were also used in this study. The global Digital Elevation Model (DEM) data (GTOPO30) with a resolution of thirty arcseconds were downloaded from the website of the US Geological Survey (USGS; https://lta.cr.usgs.gov/GTOPO30). The soil texture information was obtained from the Food and Agriculture Organization dataset [58]. The vegetation data were taken from the global vegetation classifications database developed by the University of Maryland, which has 14 types of land use/cover with a spatial resolution of 1 km × 1 km [59].

3. Methodology

3.1. Verification Metrics

To quantitatively compare GPM-era satellite precipitation products, several widely used statistical metrics were applied in this study. The formulas and perfect values of these metrics are listed in Table 2. The Pearson correlation coefficient (CC) describes the agreement between the satellite precipitation and gauge observations. Mean error (ME) simply scales the difference between the satellite estimates and the reference. The root-mean-squared error (RMSE) corresponds to the square root of the average of the squared differences between the estimates and the observed rainfall, and was used to measure the average error magnitude, while relative bias (RB) and relative root-mean-squared error (RRMSE) were used to calculate the systematic and random components of the error in the satellite precipitation products. For the contingency table metrics, following Ebert et al. [19], we adopted four skill scores, namely probability of detection (POD), false alarm rate (FAR), frequency bias index (FBI), and equitable threat score (ETS), to examine the detection capabilities in rain events of satellite estimates. The ETS is

commonly used as an overall skill measure by the numerical weather prediction community, and the POD, FAR, and FBI present complementary information about misses, false alarms, and bias.

Table 2. List of the validation statistical metrics for evaluating satellite precipitation products.

Statistic Index	Formula	Unit	Perfect Value
Correlation coefficient (CC)	$CC = \frac{\sum_{i=1}^{n}(G_i-\overline{G})(S_i-\overline{S})}{\sqrt{\sum_{i=1}^{n}(G_i-\overline{G})^2}\cdot\sqrt{\sum_{i=1}^{n}(S_i-\overline{S})^2}}$	-	1
Root-mean-squared error (RMSE)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i-G_i)^2}$	mm	0
Mean error (ME)	$ME = \frac{1}{n}\sum_{i=1}^{n}(S_i-G_i)$	mm	0
Relative root-mean-squared error (RRMSE)	$RRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i-G_i)^2}}{\frac{1}{n}\sum_{i=1}^{n}G_i}$	%	0
Relative bias (RB)	$RB = \frac{\sum_{i=1}^{n}(S_i-G_i)}{\sum_{i=1}^{n}G_i}$	%	0
Probability of detection (POD)	$POD = \frac{H}{H+M}$	-	1
False alarm ratio (FAR)	$FAR = \frac{F}{H+F}$	-	0
Frequency bias index (FBI)	$FBI = \frac{H+F}{H+M}$	-	1
Equitable Threat Score (ETS)	$ETS = \frac{H-H_e}{H+M+F-H_e}$, $H_e = \frac{(H+M)(H+F)}{n}$	-	1
Nash–Sutcliffe coefficient of efficiency (NSE)	$NSE = 1-\frac{\sum_{i=1}^{n}(Q_{obs,i}-Q_{sim,i})^2}{\sum_{i=1}^{n}(Q_{obs,i}-\overline{Q_{obs,i}})^2}$	-	1

Notation: n, number of samples; hit (H, observed rain correctly detected); miss (M, observed rain not detected); false (F, rain detected but not observed); null (N, no rain observed nor detected); S_i, satellite precipitation; G_i, gauged observation; $Q_{obs,i}$, observed streamflow; $Q_{sim,i}$, simulated streamflow; $\overline{Q_{obs,i}}$, mean value of observed streamflow.

3.2. Hydrological Model

The hydrological model used in this study was the Variable Infiltration Capacity (VIC) model [60,61], which is a grid-based distributed hydrological model maintained by the University of Washington. The distinguishing characteristic of the VIC is that it incorporates land-cover vegetation heterogeneity, multiple soil layers with spatially varying infiltration capacity, and base flow as a non-linear recession curve [62]. The model also considers dynamic changes of both water and energy balances over a grid mesh. For a more detailed description of the VIC model, the reader is referred to Nijssen et al. [63] and Xie et al. [64]. Presently, the VIC model is widely applied in hydrological simulations and evaluations of satellite precipitations over different basins [65–67]. Here, the VIC model was implemented at a spatial resolution of 0.1° for the upper Yellow River basin. The daily streamflow observations were used to calibrate and validate the VIC model. Three statistical indices (Nash–Sutcliffe efficiency (NSE), RB and RMSE) were used to assess the result of hydrologic simulation validation (Table 2).

4. Results and Discussion

4.1. Rainfall Characteristics of the TP

Figure 2 shows the spatial distributions of three-year daily mean precipitation for CGDPA and four GPM-era precipitation products over the TP. Generally, all precipitation sets shared similar spatial distribution: the mean precipitation decreased from the southeast to the northwest. One can see that a large amount of precipitation (more than 5 mm/day) concentrates over the southern region of the TP. The reason for this is the effect of the Himalayas, which intercept the moisture from the Indian Ocean monsoon and induce rainfall. In contrast, less precipitation is observed in most of the west and north, where the Westerlies do not prevail and the Indian monsoon is relatively weaker [68]. Compared with the CGDPA, a pronounced difference can be found between the IMERG and GSMaP precipitation products. IMERG-UC significantly underestimated the TP's precipitation and the GSMaP-MVK clearly overestimated the precipitation. However, the precipitation estimates of both gauge-adjusted products

were more consistent with the ground measurements than the satellite-only products. After gauge correction, the under- and overestimations of satellite-only products were mitigated, especially in the eastern TP with more rain gauges. Therefore, the strategy of gauge adjustments using in situ measurements greatly improved the product accuracy.

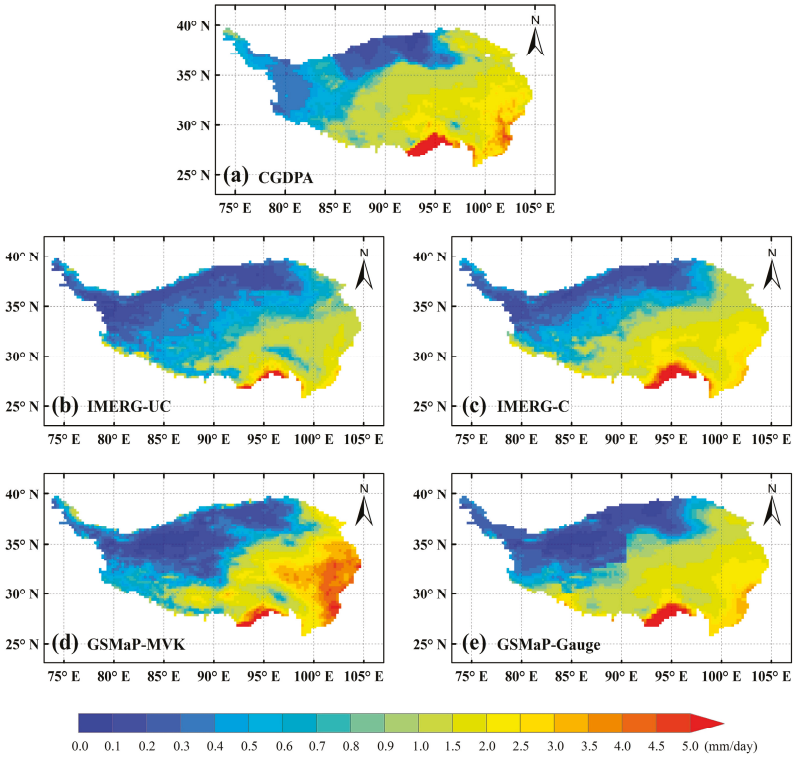


Figure 2. Spatial distributions of mean annual precipitation for (**a**) CGDPA, (**b**) IMERG-UC, (**c**) IMERG-C, (**d**) GSMaP-MVK, and (**e**) GSMaP-Gauge over the TP during the period of April 2014–March 2017.

Next, we displayed the rain rate distribution of daily precipitation amount and the daily number of precipitation events over the TP. As shown in Figure 3, for the intensity distribution from the CGDPA, the distribution of the precipitating amount over the TP presents a single-peak pattern (approximately at a rain rate of 14 mm/d), and the most precipitating days occurred in the precipitation intensity across the range of 2–8 mm·day^{-1}. The satellite precipitation products exhibit similar distribution patterns with CGDPA in terms of precipitation rates for both the occurrence and the volume of precipitation, but there are some differences in the intensity distribution curves. For example, at the light rain range (0.25–2 mm·day^{-1}), all the satellite datasets detected more precipitation events than the reference datasets (Figure 3b). Consequently, they contributed to the precipitation volumes, with the satellite-based estimates having more precipitation than the ground measurements at the light rain range.

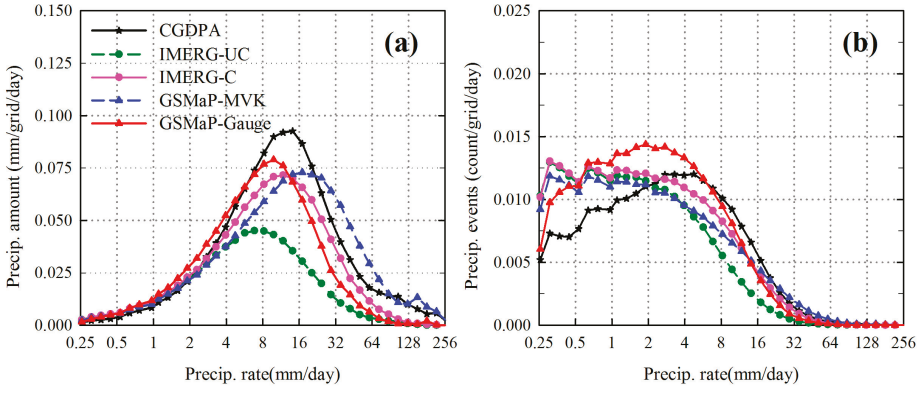


Figure 3. Intensity distribution of (**a**) daily precipitation amount (mm·day^{-1}) and (**b**) precipitation events (count per day) from the four satellite precipitation estimates at the TP. The logarithmic scale was used to bin the precipitation rates.

4.2. Statistical Performance of Satellite Precipitation Estimates

In this section, we evaluated the satellite precipitation products against gridded gauge-based precipitation products over the period from April 2014 to March 2017. In order to ensure a more accurate comparison, only grid pixels with at least one gauge (132 grids) were taken to calculate the statistical metrics.

Figure 4 shows the scatterplots of daily IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge data versus CGDPA for the selected grids. The evaluation metrics are also given in the figure. For the contingency table statistics (i.e., POD, FAR, FB, and ETS), a common threshold of 1.0 mm·day^{-1} was used to determine the rain/no rain event, as suggested by many previous studies [7,69–71]. Generally, among the four satellite precipitation products, GSMaP-Gauge exhibited the best performance with the highest CC of 0.77, while GSMaP-MVK had the worst performance with a poor CC of 0.52. The IMERG-UC and IMERG-C products showed middle performance, with CC values of 0.67 and 0.70, respectively. For the satellite-only products, we can see that IMERG-UC significantly underestimated the precipitation by about −39.32%, and GSMaP-MVK overestimated the precipitation with an RB value of 26.11%. After the gauge calibration, the RB had a downward trend and the scatter points were clustered more closely to the 1:1 line than those of satellite-only estimates. Consequently, both gauge-adjusted products only showed slight underestimation relative to reference observations (Figure 4b,d). In terms of the contingency table statistics, when compared to satellite-only products, the gauge-adjusted products also had better performance (with higher POD and ETS values). The gauge-adjusted products had more detected events compared to satellite-only products. However, it is worth noting that the FAR increased from IMERG-UC to IMERG-C. We argue that the calibration scheme of IMERG resulted in an increase in the number of false events over the TP. Thus, it contributed to the observed increases in FAR values. Correspondingly, the RMSE and RRMSE of IMERG-C did not improve over that of IMERG-UC.

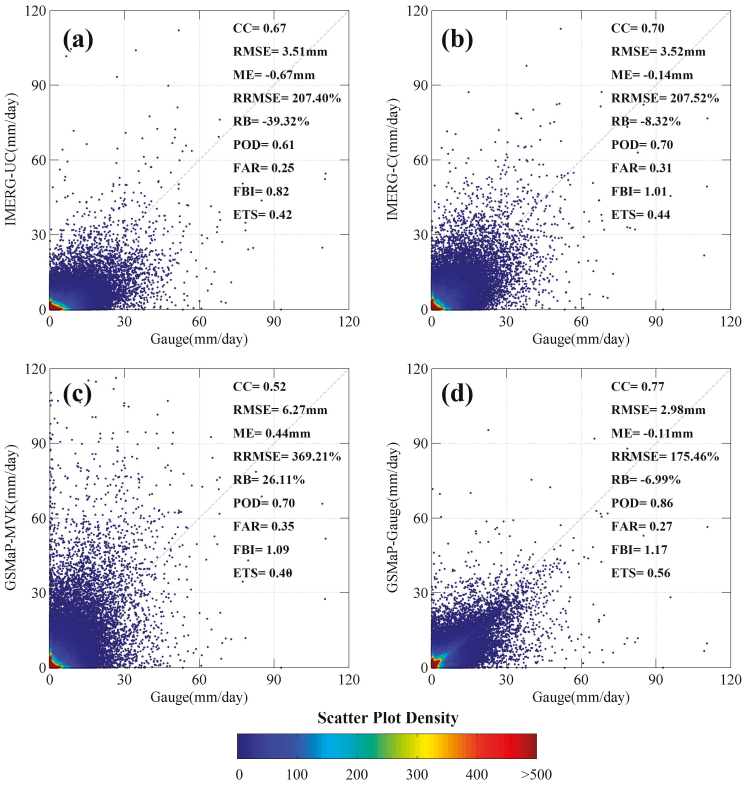


Figure 4. Scatterplots of daily precipitation for (**a**) IMERG-UC; (**b**) IMERG-C; (**c**) GSMaP-MVK; (**d**) and GSMaP-Gauge against CGDPA from the 132 selected 0.25° grid boxes over the TP.

To investigate the spatial distributions of error metrics, CC, RMSE, and RB were computed from the four satellite precipitation products, as shown in Figure 5. In general, the CC values of all products were good over most regions of TP. Spatially, higher CC values are observed in the east TP compared to the west. This pattern of CC is attributed to the limitations of retrieval of satellite precipitation in mountainous and high-elevation regions [19,66]. We also note that the GSMaP-Gauge showed best correspondence with gauge measurements, with larger correlation and smaller error (Figure 5j–l), which is consistent with the above statistical results. Interestingly, the spatial distributions of error metrics could explain the phenomenon of the RMSE of IMERG-UC being lower than that of GSMaP-MVK, but the bias performance of IMERG-UC was worse than that of GSMaP-MVK. The reason is that the positive and negative biases could cancel each other out. As shown in Figure 5c, IMERG-UC underestimated reference precipitation over almost all of TP, while GSMaP-MVK showed overestimation in the east and underestimation in the west (Figure 5i). Thus, for GSMaP-MVK, the magnitude of the total RB was reduced. However, focusing on the spatial distribution of RMSE, we can see that GSMaP-MVK still demonstrated the largest error among the four satellite precipitation products.

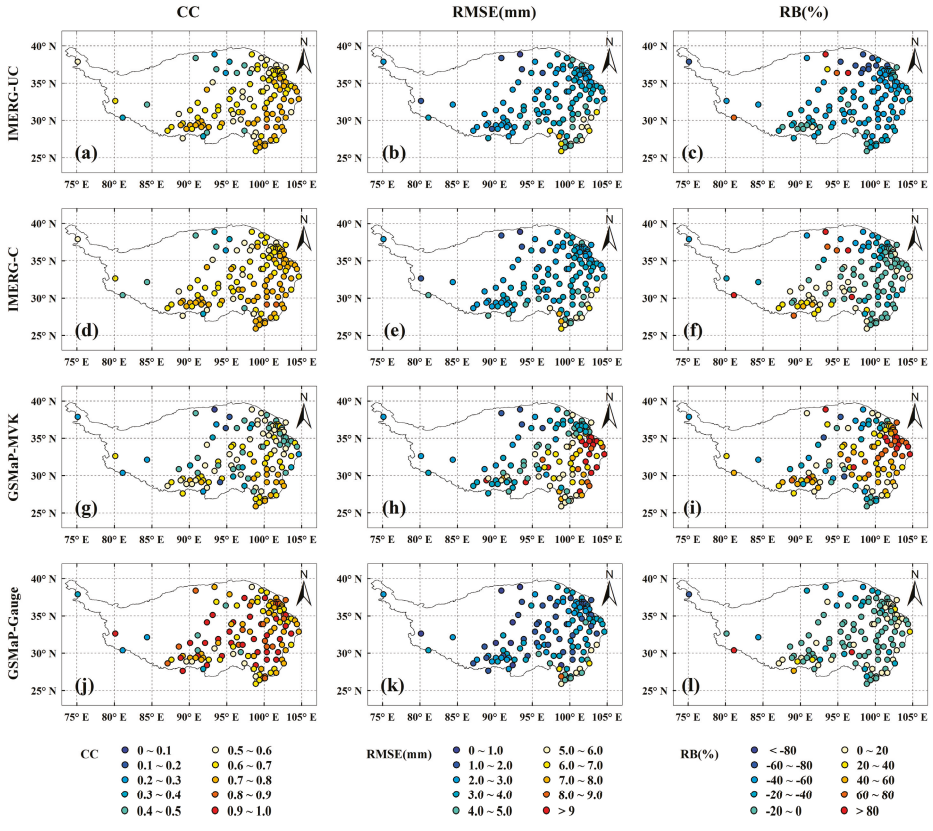


Figure 5. Spatial distribution of (**a**,**d**,**g**,**j**) correlation coefficient (CC); (**b**,**e**,**h**,**k**) root-mean-squared error (RMSE); and (**c**,**f**,**i**,**l**) relative bias (RB) between the four satellite precipitation estimates and CGDPA at the 132 grid central points.

Figure 6 shows the temporal variations of averaged spatial precipitation and statistics for the selected grid boxes. Table 3 lists the statistical summary of seasonal comparisons including spring (March–May), summer (June–August), autumn (September–November), and winter (December–February) by computing at the daily scale. Overall, the patterns of monthly mean precipitation for all satellite precipitation products exhibited similar fluctuations with gauge observations. Precipitation in the summer is the main water source over the TP, while precipitation in the winter only contributes a minor part of annual precipitation. The performance of satellite precipitation productions showed distinct seasonal variations. The statistical indices performed better in the summer than the other three seasons with high correlation, low relative error, and better detection for rain events (see Table 3). For instance, the CC value of the four satellite precipitations ranged from 0.59 to 0.76 during the summer, while a lower CC occurred in the winter. We also note that the RMSE was larger during the summer months than that of the winter season (Table 3 and Figure 6c). This is because RMSE could be affected by large precipitation bias. The RB results indicate that GSMaP-MVK overestimated the reference precipitation, except during the winter, and IMERG-UC, IMERG-C, and GSMaP-Gauge underestimated the gauge observations in all seasons. Similar to the results of Duan et al. [50] and Ning et al. [72], satellite precipitation products showed higher error and poor capability of rainfall detection in the winter months. During the winter season, although the DPR

improved the skill of snowfall observations, satellite precipitation products still showed unsatisfactory performance over the TP. This can be attributed to the limitation of passive microwave retrievals and IR information at cold or snow-covered background surfaces [22], suggesting that the current GPM-era estimates still have room for improvement in the winter.

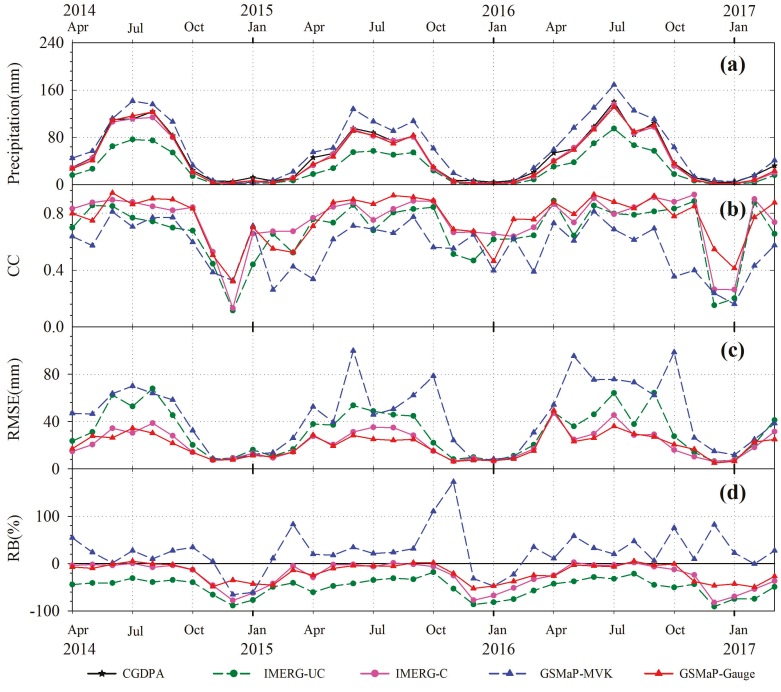


Figure 6. (**a**) Average monthly precipitation time series and monthly variations of statistical indices: (**b**) CC, (**c**) RMSE, and (**d**) RB.

Table 3. Seasonal statistics of four satellite precipitation estimates (IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge) against ground observations over the Tibetan Plateau (TP) during the study period of April 2014–March 2017.

Season	Product	CC	RMSE (mm)	ME (mm)	RRMSE (%)	RB (%)	POD	FAR	FBI	ETS
Spring	IMERG-UC	0.56	3.26	−0.62	242.87	−46.38	0.51	0.30	0.74	0.33
	IMERG-C	0.61	3.17	−0.18	236.01	−13.60	0.64	0.38	1.03	0.36
	GSMaP-MVK	0.38	6.24	0.43	464.42	31.94	0.64	0.44	1.14	0.31
	GSMaP-Gauge	0.71	2.71	−0.21	201.69	−15.53	0.82	0.30	1.17	0.52
Summer	IMERG-UC	0.67	5.15	−1.26	139.24	−34.02	0.72	0.22	0.93	0.39
	IMERG-C	0.68	5.27	−0.10	142.52	−2.74	0.80	0.26	1.07	0.40
	GSMaP-MVK	0.59	8.01	0.86	216.56	23.16	0.78	0.27	1.08	0.37
	GSMaP-Gauge	0.76	4.36	−0.08	117.87	−2.03	0.91	0.24	1.20	0.50
Autumn	IMERG-UC	0.66	3.18	−0.57	218.98	−39.09	0.60	0.25	0.81	0.41
	IMERG-C	0.71	3.12	−0.11	214.77	−7.44	0.69	0.31	1.01	0.43
	GSMaP-MVK	0.44	6.96	0.52	478.63	35.58	0.70	0.37	1.12	0.39
	GSMaP-Gauge	0.77	2.67	−0.06	183.5	−4.41	0.87	0.27	1.19	0.57
Winter	IMERG-UC	0.54	1.45	−0.21	522.96	−76.56	0.13	0.43	0.23	0.11
	IMERG-C	0.58	1.39	−0.17	500.66	−62.17	0.19	0.54	0.42	0.14
	GSMaP-MVK	0.30	1.95	−0.05	701.76	−18.30	0.29	0.66	0.85	0.16
	GSMaP-Gauge	0.58	1.40	−0.13	501.71	−45.07	0.52	0.38	0.84	0.38

For analyzing the error characteristics of different precipitation event, following the error decomposition approach proposed by Tian et al. [69], the total bias can be decomposed into different parts: hit bias, bias due to rainfall misses, bias due to false detections, and bias with selected threshold. As shown in Figure 7, hit bias and total bias share considerable similarities in their spatial distributions, suggesting that hit bias is the dominant component of total bias. Considering that the negative bias with missed precipitation and positive bias with false precipitation have opposite signs, they can offset each other, resulting in a smaller total bias of satellite precipitation. The lighter precipitation (<1 mm/day) that we considered unreliable for either gauge data or satellite measurements contributed to only a small part of total bias and can be ignored (Table 4). In addition, the error components also showed seasonal dependence. Generally speaking, it is apparent that the values of total bias are lower in the summer than in the winter. Particularly, in the winter, higher miss bias was found, and missed precipitation was the dominant source of errors. This phenomenon indicates that the satellite estimates miss a lot of precipitation events in winter, and confirms our aforementioned speculation: the GPM-era satellite precipitation products still exhibit some deficiencies for detecting snowing events.

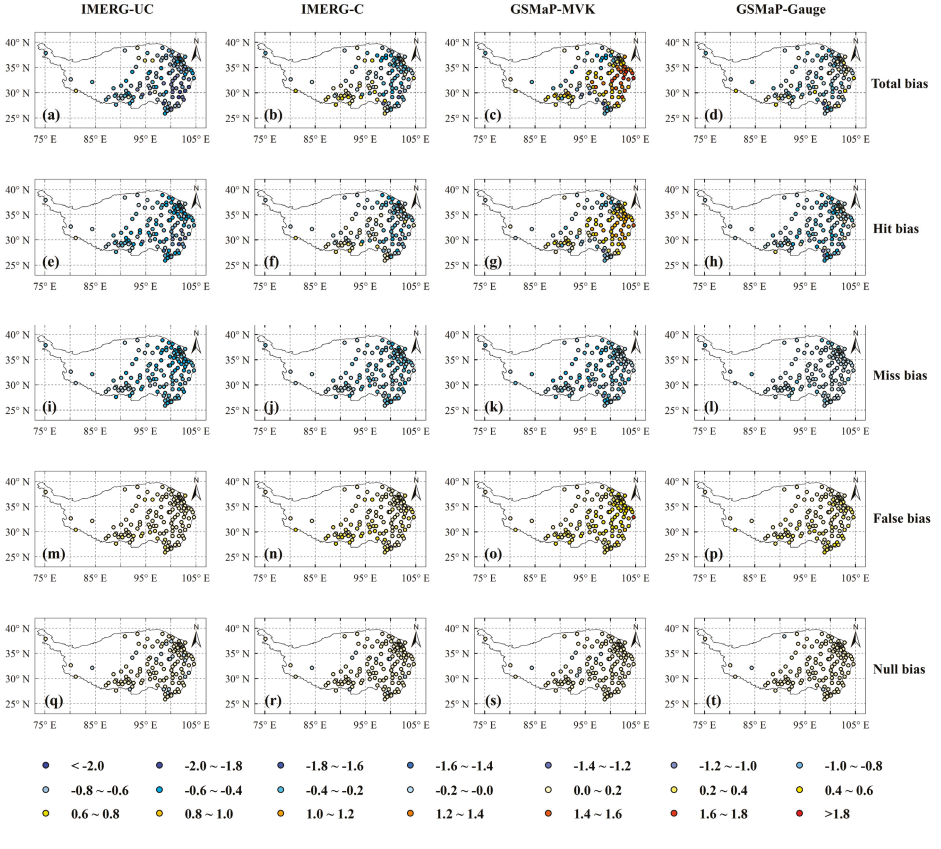


Figure 7. Spatial patterns of the error components of satellite precipitation estimates against CGDPA at the 132 grid central points (mm/day): total bias (**first row**), hit bias (**second row**), missed precipitation (**third row**), false precipitation (**fourth row**), and bias with selected threshold (**fifth row**).

Table 4. Seasonally averaged error components as percentages of total observed precipitation.

Season	Product	H (%)	−M (%)	F (%)	N (%)	E = H − M + F + N (%)
Spring	IMERG-UC	−27.84	−28.62	8.07	2.04	−46.35
	IMERG-C	−13.73	−19.75	16.66	3.24	−13.58
	GSMaP-MVK	18.94	−22.63	33.22	2.41	31.94
	GSMaP-Gauge	−21.52	−8.31	11.14	3.16	−15.53
Summer	IMERG-UC	−28.70	−12.31	6.36	0.62	−34.03
	IMERG-C	−5.89	−8.20	10.59	0.74	−2.76
	GSMaP-MVK	18.45	−10.12	13.99	0.85	23.17
	GSMaP-Gauge	−9.47	−3.26	9.49	1.20	−2.04
Autumn	IMERG-UC	−26.67	−20.41	6.69	1.30	−39.09
	IMERG-C	−6.72	−14.80	12.22	1.88	−7.42
	GSMaP-MVK	19.75	−15.31	28.57	2.57	35.58
	GSMaP-Gauge	−11.85	−5.85	11.28	2.00	−4.42
Winter	IMERG-UC	−12.94	−67.39	4.42	−0.65	−76.56
	IMERG-C	−13.16	−60.48	9.31	2.19	−62.14
	GSMaP-MVK	−6.04	−55.34	37.97	5.07	−18.34
	GSMaP-Gauge	−32.33	−28.44	11.40	4.31	−45.06
All Season	IMERG-UC	−27.46	−19.59	6.71	1.01	−39.33
	IMERG-C	−7.96	−14.06	12.14	1.56	−8.32
	GSMaP-MVK	17.89	−15.59	22.09	1.72	26.11
	GSMaP-Gauge	−13.35	−5.84	10.30	1.90	−6.99

Figure 8 displays the error characteristics of satellite precipitation estimates with rain rate. Obviously, all precipitation products showed a similar variation of error, with overestimations for light rain and underestimations for heavy rain (Figure 8a), which is a common error feature of satellite-based retrievals, as documented in previous studies [73–75]. This error feature of rain-rate dependency is important for meteorological and hydrological applications, especially for typhoon monitoring and flood forecast, which are sensitive to higher rain rates [76,77]. In terms of RRMSE, higher values were found at low rain rates compared to at moderate–high rain rates (Figure 8b), indicating that current satellite precipitation products need to continue improving the performance at low rain rates. On the other hand, from the results of Figure 8, we can see how the calibration scheme works in precipitation estimates. It is seen that the IMERG-C elevated the precipitation estimates and GSMaP-Gauge decreased the precipitation values compared to their corresponding uncalibrated precipitation products. This gauge-calibration effectively reduced total bias while making things worse in some cases. For example, IMERG-UC overestimated gauge observations at lower rain rates; however, the bias calibration using GPCC gauge data elevated the precipitation estimates further augmenting the overestimation at lower rain rates. Another case involved the GSMaP-Gauge showing large negative hit bias than GSMaP-MVK in the winter season (Table 4). Thus, it seems important to calibrate satellite-based precipitation estimates at difference rain rates or seasons in the future.

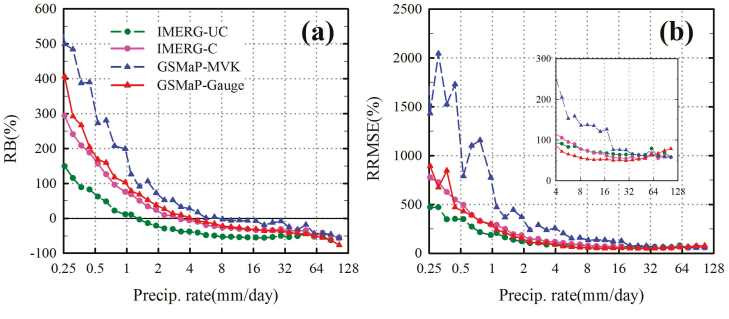


Figure 8. The relationship between the validation indices and precipitation rate: (**a**) RB, and (**b**) RRMSE.

4.3. Hydrological Evaluation of Satellite Precipitation Estimates

In the previous section, we compared the GPM-era satellite precipitation products against the rain gauge observations; the next step was to evaluate the hydrological utility of these precipitation datasets. In this section, since the streamflow data after 2015 were not available at the Tangnaihai hydrological station, the hydrological evaluation of four satellite precipitation estimates was performed for the whole year of 2015. We also calculated statistical indices of precipitation estimates over the upper Yellow River basin in 2015 (Table 5). By analyzing this indices, we can conclude that the error characteristics of satellite precipitation during 2005 are consistent with prior comparison results. For example, the CC values of daily IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge estimates were 0.57, 0.61, 0.52, and 0.75 over the upper Yellow River basin during 2005, and 0.67, 0.70, 0.52, and 0.77 over the TP during the periods of April 2014 to March 2017, respectively. For the basin-scale evaluation, statistical values with the basin-averaged data were better than those with the grid-scale evaluation. This is expected because random errors would decrease with spatial scale averaging. Thus, both grid-scale and basin-scale analysis confirm that the performance of satellite precipitation in the upper Yellow River basin was similar to that on the TP.

Table 5. Statistical summary of the precipitation products at grid and basin scales for the year of 2015 in the upper Yellow River basin.

	Product	CC	RMSE (mm)	RB (%)
Grid scale	IMERG-UC	0.57	3.10	−46.65
	IMERG-C	0.61	3.30	−12.05
	GSMaP-MVK	0.52	5.78	59.25
	GSMaP-Gauge	0.75	2.41	−3.19
Basin scale	IMERG-UC	0.74	1.65	−46.07
	IMERG-C	0.79	1.60	−12.19
	GSMaP-MVK	0.74	2.70	48.01
	GSMaP-Gauge	0.93	0.82	−4.90

Next, the VIC model was calibrated and validated with observed precipitation and streamflow for the periods of 2009–2011 and 2012–2014 over the upper Yellow River basin. Figure 9 shows the CGDPA-simulated and observed streamflow at the daily scale. Comparing the observed and simulated streamflow, the values of NSE and RB were 0.61% and −2.56% during the calibration period, and NSE increased to 0.73 and RB of 0.96% in the validation period. It can be seen that the simulated streamflow generally agrees with observations very well, although overestimation and underestimation of peak floods existed in some cases. Improved results were obtained at the monthly scale for both the calibration and validation periods (NSE of 0.88 and 0.87, respectively).

After the model was benchmarked by the in situ data, the VIC model was then driven by gauge- and satellite-based precipitation datasets for the period from 1 January 2015 to 31 December 2015, without any further adjustment of parameters. The simulated and observed hydrographs are shown in Figure 10, and the statistical comparisons are summarized in Table 6. As shown, the CGDPA had a worse performance in 2015, with NSE of 0.41 and a runoff overestimation of 26.39%. The observed mean daily discharge from 2009 to 2014 was 711.61 m^3/s, whereas that in 2015 was 480.82 m^3/s. Differences in hydrological features during the two periods may potentially influence the simulation performance. Using the same parameters enabled us to compare the performance of simulated streamflow from different precipitation inputs. For the GPM-era satellite precipitation products, the GSMaP-Gauge showed the best performance in the streamflow simulation; the IMERG-C took second place; and the two purely satellite-derived estimates demonstrated poor performance due to the large precipitation bias, especially for GSMaP-MVK with 151.97% runoff overestimation at a daily scale. Interestingly, the simulated streamflow with GSMaP-Gauge inputs had lightly better performance than the CGDPA (e.g., 0.53 versus 0.41 for NSE). We considered that gauge corrections involved in the GSMaP-Gauge

products remarkably improved the skill of streamflow simulation. For the monthly comparisons, GSMaP-Gauge performed the best again, while the NSE value of IMERG-C reached 0.63. However, the satellite-only products still had unsatisfactory performance with negative NSEs, suggesting they have low hydrological utility for this region.

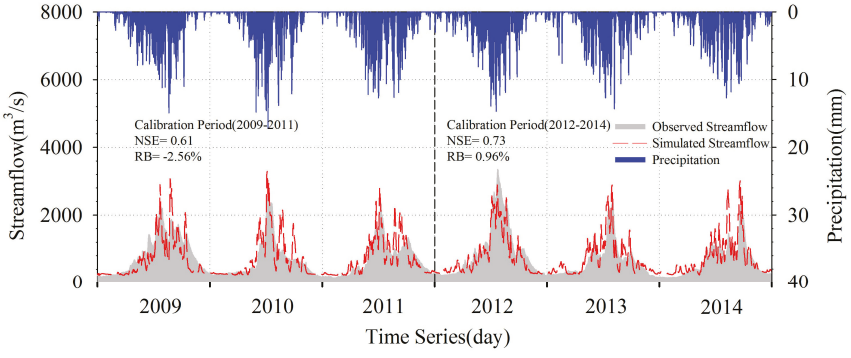


Figure 9. Observed and Variable Infiltration Capacity (VIC) model simulated streamflow with the CGDPA precipitation for the calibration period (2009–2011), and validation period (2012–2014) over the upper Yellow River basin.

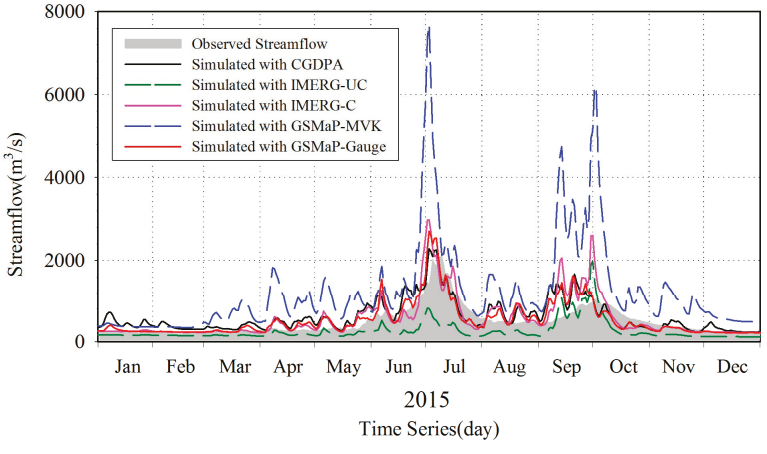


Figure 10. Daily observed and simulated streamflow with gauge benchmarked parameters over the upper Yellow River basin.

The simulation accuracy could be improved if the hydrology model was calibrated with different precipitation inputs. Subsequently, we recalibrated the model parameters using each satellite precipitation dataset during 2015. This scenario is also an alternative strategy for hydrological applications in ungauged basins where only satellite precipitation estimates are available [78,79]. As shown in Figure 11, the simulation performances were effectively improved after the model was recalibrated. For example, the daily NSE of IMERG-C significantly increased from 0.18 to 0.63. Furthermore, the RMSE also significantly decreased for all satellite products. As summarized in Table 6, simulations of IMERG-C and GSMaP-Gauge products had good statistical agreement with observed streamflow at daily and monthly scale. However, the NSE values of both satellite-only products was still below zero at the daily scale, further confirming that the undesirable hydrological utility of these

two satellite-only products is mainly due to the unreliable precipitation estimates. The errors existing in these two precipitation datasets was propagated to simulated streamflow, and could not be removed upon model-parameter recalibration. Generally, the recalibration of the model parameters effectively improved the hydrological potential of satellite precipitation, especially for precipitation products with small errors; however, this recalibration approach should be taken with a grain of salt because it may result in unrealistic parameter values in some cases [80,81].

Table 6. Comparison of daily and monthly observed and simulated streamflow when the Variable Infiltration Capacity (VIC) model was forced by the gauge- and satellite-based precipitation datasets in 2015.

Time Scales	Precipitation Products	Benchmarking Calibration			Product-Specific Calibration		
		NSE	RB (%)	RMSE (m^3/s)	NSE	RB (%)	RMSE (m^3/s)
Daily	CGDPA	0.41	26.39	274.05	0.63	0.55	217.11
	IMERG-UC	−0.12	−46.00	377.46	−0.08	−18.96	370.33
	IMERG-C	0.18	17.97	323.77	0.62	−7.26	220.87
	GSMaP-MVK	−9.85	151.97	1175.70	−2.93	66.73	707.77
	GSMaP-Gauge	0.53	11.64	243.91	0.67	0.13	205.56
Monthly	CGDPA	0.53	26.48	214.93	0.70	0.62	168.61
	IMERG-UC	−0.17	−45.96	333.21	0.15	−18.90	286.48
	IMERG-C	0.63	18.05	191.45	0.76	−7.19	150.74
	GSMaP-MVK	−6.88	152.16	876.17	−1.85	66.85	518.21
	GSMaP-Gauge	0.74	11.73	160.89	0.79	0.21	140.43

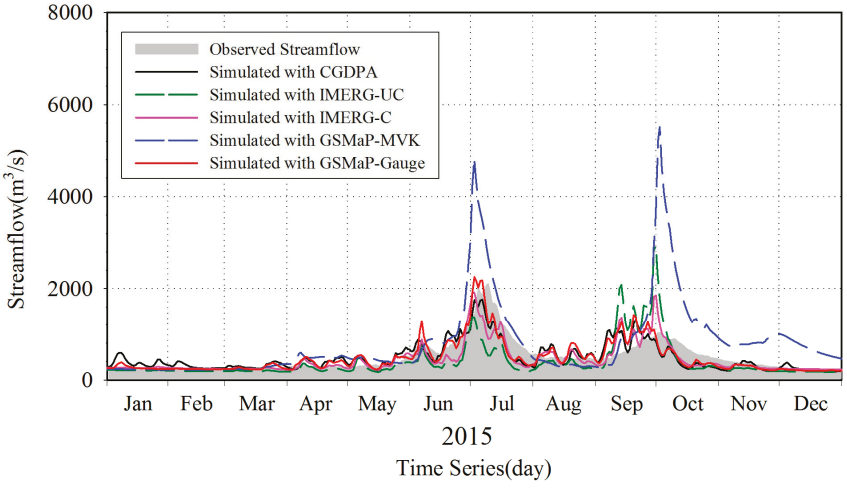


Figure 11. Comparison of VIC simulated streamflow with recalibrated parameters using product-specific inputs.

5. Conclusions and Recommendations

After four years of operation of the GPM, it was proven to be a good successor to the TRMM, offering new opportunities for meteorological studies and hydrological applications, especially in mountainous areas. In this study, we investigated the performance of the latest GPM IMERG V5 and GSMaP V7 satellite precipitation products over the TP. Firstly, the statistical assessment of the accuracy of the satellite products was performed against gridded gauge-based data. Then, their hydrological simulation utilities were evaluated using the VIC model in the upper Yellow River basin.

Our study shows that IMERG and GSMaP products can appropriately capture spatial patterns of precipitation across the TP. IMERG-UC significantly underestimated the reference precipitation (−39.32%), while obvious overestimation was found in GSMaP-MVK (26.11%). After bias adjustment, the under- and overestimations were clearly reduced, with slight underestimation for IMERG-C (−8.32%) and GSMaP-Gauge (−6.99%). Among the four studied satellite estimates, GSMaP-Gauge had the best performance in nearly all statistical indices with higher correlation, lower bias, and better detection, while GSMaP-MVK had worst performance. This results suggest that the calibration procedure in GSMaP-gauge performs better than that in IMERG-C. Considering that IMERG-C ingests monthly GPCC gauge analyses and that GSMaP-Gauge is calibrated by CPC daily gauge analyses, it seems important to use daily gauge data for IMERG in the future. In fact, the GPCC started computing daily adjustments based on station meteorological data, and it will be used in a future version of IMERG [47].

The GPM sensors were designed to extend the capability for detecting light-intensity precipitation and falling snow, making IMERG perform better than TMPA, which was confirmed by previous studies [29,32,82]. However, in the present study, we found that all satellite precipitation products had a relatively lower accuracy in winter, and missed bias was the main problem during the winter season. On the other hand, higher error values were found at low rain rates compared to at moderate–high rain rates. Such results indicate that the current GPM-era satellite precipitation products still have significant room for improving the estimation of light rainfall and falling snow events.

The hydrological simulations presented here show that the gauge-adjusted products are more promising than their unadjusted counterparts. The satellite-only products have unsatisfactory performance with large runoff errors. Even when the VIC model was recalibrated using precipitation-specific inputs, IMERG-UC and GSMaP-MVK still showed little potential for streamflow simulation over the upper Yellow River basin. This result implies that the VIC hydrologic model can tolerate relatively small precipitation errors but not unphysical bias through the hydrologic integration process. Thus, if a satellite precipitation estimate has good intrinsic data quality, then precipitation-specific calibration is recommended. Clearly, the performances of satellite precipitation estimates rely on gauge corrections, but purely satellite-derived estimates are the only available source of precipitation for areas where ground observations are nonexistent. At the current stage, the climatological calibration is proposed to make the purely satellite-derived product as consistent as possible with gauge-corrected products. In the future, more PMW/IR sources and more effective precipitation algorithms are needed for improving the retrieval ability of satellite precipitation.

We hope this study here will be helpful as a reference for precipitation dataset selection in the TP, and also provide the satellite precipitation community with a better understanding of GPM-era precipitation products. Considering the finer spatiotemporal resolution and improved precision, more studies exploring the potential of IMERG and GSMaP precipitation products in sub-daily or hourly time scales will be conducted in the near future.

Author Contributions: D.L. and B.Y. designed the study and conducted the experiments; D.L. wrote the draft of the manuscript; B.Y. supervised the research and revised the manuscript.

Funding: This work was financially supported by the National Key Research and Development Program of China (2018YFA0605402) and the Major Research Plan of the National Natural Science Foundation of China (91547101). Also, this work was partially sponsored by the Natural Science Foundation of Jiangsu Province (BK20161502).

Acknowledgments: The authors are grateful to the GPM science team at NASA and JAXA for making satellite precipitation data available, and thankful to the CMA for providing observation gauge data.

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

References

1. Allen, M.R.; Ingram, W.J. Constraints on future changes in climate and the hydrologic cycle. *Nature* **2002**, *419*, 224–232. [CrossRef] [PubMed]

2. 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]
3. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116. [CrossRef]
4. 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]
5. 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]
6. Qiu, J. The third pole. *Nature* **2008**, *454*, 393–396. [CrossRef] [PubMed]
7. Tong, K.; Su, F.; Yang, D.; Hao, Z. Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *J. Hydrol.* **2014**, *519*, 423–437. [CrossRef]
8. Hossain, F.; Lettenmaier, D.P. Flood prediction in the future: Recognizing hydrologic issues in anticipation of the Global Precipitation Measurement mission. *Water Resour. Res.* **2006**, *42*, W11301. [CrossRef]
9. Katsanos, D.; Retalis, A.; Tymvios, F.; Michaelides, S. Analysis of precipitation extremes based on satellite (CHIRPS) and in situ dataset over Cyprus. *Nat. Hazards* **2016**, *83*, 53–63. [CrossRef]
10. 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]
11. 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]
12. 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]
13. Turk, F.J.; Miller, S.D. Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Rem. Sens.* **2005**, *43*, 1059–1069. [CrossRef]
14. 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]
15. 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]
16. Hong, Y.; Adler, R.F.; Negri, A.; Huffman, G.J. Flood and landslide applications of near real-time satellite rainfall products. *Nat. Hazards* **2007**, *43*, 285–294. [CrossRef]
17. Wu, H.; Adler, R.F.; Tian, Y.; Huffman, G.J.; Li, H.; Wang, J. Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model. *Water Resour. Res.* **2014**, *50*, 2693–2717. [CrossRef]
18. Yang, N.; Zhang, K.; Hong, Y.; Zhao, Q.; Huang, Q.; Xu, Y.; Xue, X.; Chen, S. Evaluation of the TRMM multisatellite precipitation analysis and its applicability in supporting reservoir operation and water resources management in Hanjiang basin, China. *J. Hydrol.* **2017**, *549*, 313–325. [CrossRef]
19. 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]
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. Tian, Y.; Peters-Lidard, C.D. A global map of uncertainties in satellite-based precipitation measurements. *Geophys. Res. Lett.* **2010**, *37*, L24407. [CrossRef]
22. 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]

23. 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]
24. 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]
25. Skofronick-Jackson, G.; Berg, W.; Kidd, C.; Kirschbaum, D.B.; Petersen, W.A.; Huffman, G.J.; Takayabu, Y.N. Global Precipitation Measurement (GPM): Unified Precipitation Estimation from Space. In *Remote Sensing of Clouds and Precipitation*; Springer International Publishing AG: Cham, Switzerland, 2018; pp. 175–193, ISBN 978-3-319-72582-6.
26. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC Code. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc_180207.pdf (accessed on 6 September 2018).
27. Skofronick-Jackson, G.; Huffman, G.; Petersen, W. Three Years of the Global Precipitation Measurement (GPM) Mission. Available online: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20180000664.pdf (accessed on 6 September 2018).
28. Gaona, M.R.; Overeem, A.; Leijnse, H.; Uijlenhoet, R. First-year evaluation of GPM rainfall over the Netherlands: IMERG day 1 final run (V03D). *J. Hydrometeorol.* **2016**, *17*, 2799–2814. [CrossRef]
29. 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]
30. Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Liu, Z.; Norouzi, H.; Pai, D.S. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.* **2016**, *556*, 865–876. [CrossRef]
31. 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. Hydrometeor.* **2016**, *17*, 2875–2882. [CrossRef]
32. 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]
33. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Day 1 IMERG Final Run Release Notes. NASA/GSFC: Greenbelt, MD, USA. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_FinalRun_Day1_release_notes.pdf (accessed on 6 September 2018).
34. 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]
35. Oliveira, R.; Maggioni, V.; Vila, D.; Porcacchia, L. Using Satellite Error Modeling to Improve GPM-Level 3 Rainfall Estimates over the Central Amazon Region. *Remote Sens.* **2018**, *10*, 336. [CrossRef]
36. 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]
37. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]
38. Asong, Z.E.; Razavi, S.; Wheater, H.S.; Wong, J.S. Evaluation of integrated multisatellite retrievals for GPM (IMERG) over southern Canada against ground precipitation observations: A preliminary assessment. *J. Hydrometeorol.* **2017**, *18*, 1033–1050. [CrossRef]
39. 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]
40. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Stocker, E.F.; Tan, J. *V05 IMERG Final Run Release Notes*; NASA Goddard Earth Sciences Data and Information Services Center: Greenbelt, MD, USA, 2018. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_FinalRun_V05_release_notes-rev3.pdf (accessed on 6 September 2018).

41. 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. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 2712–2715. [CrossRef]
42. Immerzeel, W.W.; Van Beek, L.P.H.; Bierkens, M.F.P. Climate change will affect the Asian water towers. *Science* **2010**, *328*, 1382–1385. [CrossRef]
43. Guo, D.; Wang, H. The significant climate warming in the northern Tibetan Plateau and its possible causes. *Int. J. Climatol.* **2012**, *32*, 1775–1781. [CrossRef]
44. Tong, K.; Su, F.; Yang, D.; Zhang, L.; Hao, Z. Tibetan Plateau precipitation as depicted by gauge observations, reanalyses and satellite retrievals. *Int. J. Climatol.* **2014**, *34*, 265–285. [CrossRef]
45. Rees, H.G.; Collins, D.N. Regional differences in response of flow in glacier-fed Himalayan rivers to climatic warming. *Hydrol. Process.* **2006**, *20*, 1493–1517. [CrossRef]
46. Cui, X.; Graf, H.F. Recent land cover changes on the Tibetan Plateau: A review. *Clim. Chang.* **2009**, *94*, 47–61. [CrossRef]
47. 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 Integrated MultisatellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., Version 5.2. Available online: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2_0.pdf (accessed on 6 September 2018).
48. 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]
49. 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]
50. 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]
51. Ushio, T.; Sasashige, K.; Kubota, T.; Shige, S.; Okamoto, K.I.; 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]
52. 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. Hydrometeorol.* **2007**, *8*, 607–626. [CrossRef]
53. 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**. [CrossRef]
54. Shen, Y.; Xiong, A.; Wang, Y.; Xie, P. Performance of high-resolution satellite precipitation products over China. *J. Geophys. Res. Atmos.* **2010**, *115*. [CrossRef]
55. 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]
56. 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]
57. 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]
58. FAO. *Food and Agriculture Association: Digital Soil Map of the World and Derived Soil Properties, Land and Water Digital Media Series*; CD-ROM: Rome, Italy, 2003.
59. Hansen, M.C.; DeFries, R.S.; Townshend, J.R.G.; Sohlberg, R. Global land cover classification at 1 km resolution using a classification tree approach. *Int. J. Remote Sens.* **2000**, *21*, 1331–1364. [CrossRef]
60. Liang, X.; Lettenmaier, D.P.; Wood, E.F.; Burges, S.J. A simple hydrologically based model of land surface water and energy fluxes for GSMs. *J. Geophys. Res. Atmos.* **1994**, *99*, 14415–14428. [CrossRef]
61. Liang, X.; Wood, E.F.; Lettenmaier, D.P. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Glob. Planet. Chang.* **1996**, *13*, 195–206. [CrossRef]

62. Gao, H.; Tang, Q.; Shi, X.; Zhu, C.; Bohn, T.; Su, F.; Sheffield, J.; Pan, M.; Lettenmaier, D.; Wood, E.F. Water Budget Record from Variable Infiltration Capacity (VIC) Model Algorithm Theoretical Basis Document, Algorithm Theoretical Basis Document for Terrestrial Water Cycle Data Records. Available online: http://dynamo.hydro.washington.edu/SurfaceWaterGroup/Publications/Water_Cycle_MEaSUREs_ATBD_VICmodel_submit.doc (accessed on 6 September 2018).
63. Nijssen, B.; Lettenmaier, D.P.; Liang, X.; Wetzel, S.W.; Wood, E.F. Streamflow simulation for continental-scale river basins. *Water Resour. Res.* **1997**, *33*, 711–724. [CrossRef]
64. Xie, Z.; Yuan, F.; Duan, Q.; Zheng, J.; Liang, M.; Chen, F. Regional parameter estimation of the VIC land surface model: Methodology and application to river basins in China. *J. Hydrometeorol.* **2007**, *8*, 447–468. [CrossRef]
65. Su, F.; Hong, Y.; Lettenmaier, D.P. Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata basin. *J. Hydrometeorol.* **2008**, *9*, 622–640. [CrossRef]
66. 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*, W07542. [CrossRef]
67. Sun, R.; Yuan, H.; Liu, X.; Jiang, X. Evaluation of the latest satellite–gauge precipitation products and their hydrologic applications over the Huaihe River basin. *J. Hydrol.* **2016**, *536*, 302–319. [CrossRef]
68. 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]
69. 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*, D24101. [CrossRef]
70. 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]
71. 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]
72. Ning, S.W.; Wang, J.; Jin, J.L.; Ishidaira, H. Assessment of the latest GPM-Era high-resolution satellite precipitation products by comparison with Observation Gauge Data over the Chinese Mainland. *Water* **2016**, *8*, 481. [CrossRef]
73. 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]
74. Su, J.; Lü, H.; Zhu, Y.; Wang, X.; Wei, G. Component Analysis of Errors in Four GPM-Based Precipitation Estimations over Mainland China. *Remote Sens.* **2018**, *10*, 1420. [CrossRef]
75. Wang, D.; Wang, X.; Liu, L.; Wang, D.; Huang, H.; Pan, C. Evaluation of TMPA 3B42V7, GPM IMERG and CMPA precipitation estimates in Guangdong Province, China. *Int. J. Climatol.* **2018**. [CrossRef]
76. Omranian, E.; Sharif, H.; Tavakoly, A. How well can global precipitation measurement (GPM) capture hurricanes? case study: Hurricane Harvey. *Remote Sens.* **2018**, *10*, 1150. [CrossRef]
77. 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]
78. Xue, X.; Hong, Y.; Limaye, A.S.; Gourley, J.J.; Huffman, G.J.; Khan, S.I.; Dorji, C.; Chen, S. Statistical and hydrological evaluation of TRMM-based multi-satellite precipitation analysis over the Wangchu basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J. Hydrol.* **2013**, *499*, 91–99. [CrossRef]
79. Yuan, F.; Wang, B.; Shi, C.; Cui, W.; Zhao, C.; Liu, Y.; Ren, L.; Zhang, L.; Zhu, Y.; Chen, T.; et al. Evaluation of hydrological utility of IMERG Final Run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. *J. Hydrol.* **2018**. [CrossRef]
80. Maggioni, V.; Massari, C. On the performance of satellite precipitation products in riverine flood modeling: A review. *J. Hydrol.* **2018**, *558*, 214–224. [CrossRef]

81. 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]
82. Prakash, S.; Kumar, M.R.; Mathew, S.; Venkatesan, R. How accurate are satellite estimates of precipitation over the north Indian Ocean? *Theor. Appl. Climatol.* **2017**, *134*, 467–475. [CrossRef]

remote sensing

MDPI

Article

Assessment of Ground-Reference Data and Validation of the H-SAF Precipitation Products in Brazil

Lia Martins Costa do Amaral [1,*], Stefano Barbieri [2], Daniel Vila [1], Silvia Puca [3], Gianfranco Vulpiani [3], Giulia Panegrossi [4], Thiago Biscaro [1], Paolo Sanò [4], Marco Petracca [3], Anna Cinzia Marra [4], Marielle Gosset [5] and Stefano Dietrich [4]

1 Weather Forecast Center and Climate Studies, National Institute for Space Research (CPTEC/INPE), São José dos Campos, SP 12227-010, Brazil; daniel.vila@inpe.br (D.V.); thiago.biscaro@inpe.br (T.B.)
2 CETEMPS, University of L'Aquila, 67100 L'Aquila, Italy; barbieris380@gmail.com
3 Italian Civil Protection Department, 00189 Rome, Italy; silvia.puca@protezionecivile.it (S.P.); gianfranco.vulpiani@protezionecivile.it (G.V.); Marco.Petracca@protezionecivile.it (M.P.)
4 Institute of Atmospheric Sciences and Climate (ISAC) National Research Council of Italy (CNR), 00133 Rome, Italy; giulia.panegrossi@artov.isac.cnr.it (G.P.); paolo.sano@artov.isac.cnr.it (P.S.); anna.cinzia.marra@gmail.com (A.C.M.); s.dietrich@isac.cnr.it (S.D.)
5 Institute of Research for Development (IRD), 13572 Marseille, France; marielle.gosset@ird.fr
* Correspondence: lia.meteorologia@gmail.com; Tel.: +55-12-3208-6645

Received: 13 August 2018; Accepted: 24 October 2018; Published: 5 November 2018

Abstract: The uncertainties associated with rainfall estimates comprise various measurement scales: from rain gauges and ground-based radars to the satellite rainfall retrievals. The quality of satellite rainfall products has improved significantly in recent decades; however, such algorithms require validation studies using observational rainfall data. For this reason, this study aims to apply the H-SAF consolidated radar data processing to the X-band radar used in the CHUVA campaigns and apply the well established H-SAF validation procedure to these data and verify the quality of EUMETSAT H-SAF operational passive microwave precipitation products in two regions of Brazil (Vale do Paraíba and Manaus). These products are based on two rainfall retrieval algorithms: the physically based Bayesian Cloud Dynamics and Radiation Database (CDRD algorithm) for SSMI/S sensors and the Passive microwave Neural network Precipitation Retrieval algorithm (PNPR) for cross-track scanning radiometers (AMSU-A/AMSU-B/MHS sensors) and for the ATMS sensor. These algorithms, optimized for Europe, Africa and the Southern Atlantic region, provide estimates for the MSG full disk area. Firstly, the radar data was treated with an overall quality index which includes corrections for different error sources like ground clutter, range distance, rain-induced attenuation, among others. Different polarimetric and non-polarimetric QPE algorithms have been tested and the Vulpiani algorithm (hereafter, R_{q2Vu15}) presents the best precipitation retrievals when compared with independent rain gauges. Regarding the results from satellite-based algorithms, generally, all rainfall retrievals tend to detect a larger precipitation area than the ground-based radar and overestimate intense rain rates for the Manaus region. Such behavior is related to the fact that the environmental and meteorological conditions of the Amazon region are not well represented in the algorithms. Differently, for the Vale do Paraíba region, the precipitation patterns were well detected and the estimates are in accordance with the reference as indicated by the low mean bias values.

Keywords: rain gauges; radar; quality indexes; satellite rainfall retrievals; validation

1. Introduction

The knowledge about the distribution of water around the globe is an aspect of extreme relevance for the management of natural resources. The precipitation is, within the hydrological

cycle, unanimously recognized as a central component, regulating the energy balance through the interactions of water vapor and clouds, where redistribution of latent heat occurs in the atmosphere. A detailed characterization about precipitation, its formation processes and its life cycle is essential to improve the quality of weather and climate forecasts and also to help decision-makers in their resolutions to be taken in areas affected by the rain.

Due to its inherent complexity, rainfall presents high spatial and temporal variability resulting in different regimes, all these factors combined make the tasks of observation, comprehension and prediction more challenging. The quantification of precipitation, in terms of frequency and intensity, is performed by rain gauge and meteorological radars [single and double polarization] at ground. However, these, in turn, have a sparse and uneven distribution, especially in mountainous regions, forest and over the ocean. Thus, satellite-based precipitation estimates fill these gaps and complement the rainfall observation system.

The quality of satellite rainfall products has improved significantly in recent decades, especially with the advent of satellites/missions such as the Tropical Rainfall Measuring Mission (TRMM) [1] and the Global Precipitation Measurement (GPM) [2]. In addition to the greater number of channels and radiometers available to clouds and precipitation exploitation, the improvements to computational methods and precipitation clouds modeling provided the possibility of development of several rainfall retrieval techniques (i.e, [3–5]).

In this context, the EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF) provides rainfall estimations based on infrared and microwave satellite sensors aboard polar and geostationary satellites. On each new Continuous Development Phase (CDOP), new products are released and validation processes are necessary to verify the algorithms performance both in the H-SAF area, as in extra-European regions. The validation of these satellite retrievals is performed by the H-SAF Precipitation Product Validation Group (PPVG) where a common validation methodology has been defined inside the PPVG in order to make those validation results from several institutes comparable and understandable [6,7].

From the validation perspective, the rain gauge networks can provide accurate rainfall information; however they are point measurements and are unevenly distributed around the globe. On the other hand, meteorological radars can provide far better coverage in space and time, especially considering dual polarization radars that provide information about the hydrometeors such as size, shape and variety, which allows for better understanding cloud and precipitation microphysics. Nonetheless, an important issue related to radar systems is the error sources associated with it, which makes challenging the quantitative precipitation estimation (QPE)-based solely on radar, unless these error sources are properly treated. Typically, the error sources that can be identified are: radar calibration, ground-clutter, wet-radome attenuation, rain-induced attenuation, vertical profile of reflectivity (VPR) and non-uniform beam filling. Despite this, polarimetric radars offer new variables that allow greater insight in precipitation and new ways to deal with these error sources [8].

Regarding the quantification of the characteristics of the uncertainties (random components and systematic errors) associated with satellite precipitation estimates, several efforts have been carried out around the world. One example is the CHUVA Project (Cloud processes of the main precipitation systems in Brazil: A contribution to cloud resolving modeling and to the GPM [Global Precipitation Measurement]) that aims to study the clouds and precipitation processes, through six field experiments under different precipitation regimes in Brazil. One of the major goals of the CHUVA project is to evaluate and improve the quality of satellite-based precipitation estimates from the GPM constellation [9]. More about the CHUVA project and its field campaigns can be found online in the CHUVA Project website: http://chuvaproject.cptec.inpe.br.

This study has two main goals: (i) to apply the H-SAF consolidated radar data processing to the X-band radar used in the CHUVA campaigns and; (ii) apply the consolidated H-SAF validation procedure to these data and verify the quality of H-SAF products over specific regions in Brazil. In Section 2, the study areas, radar characteristic and rain gauge distribution are described. In the same

section, the methodology for radar data treatment (quality index) and radar-based rainfall algorithms are described. A brief description of satellite algorithms is also provided. In Section 3 the results of the evaluation of rainfall radar estimates with respect to the rain gauges are presented, as well the results from the satellite verification process, where statistical and pixel by pixel evaluations are performed. Finally, a summary of the main results is presented in Section 5.

2. Materials and Methods

2.1. Study Area and Data Sources (Radar and Rain Gauges)

This study was developed based on two CHUVA field campaigns. The first experiment occurred in the Vale do Paraíba region located in the southeastern part of Brazil and the second campaign took place in Manaus city in the northern region of Brazil (centrally located in the Amazon basin). The satellite-based evaluation process considers the precipitating events with largest rain rates for both campaigns. Specifically, the case studies analyzed in Manaus took place on 15, 21, 23, 24, 25, 26 February and 2 and 8 March 2014 (8 days) while in the Vale do Paraíba campaign 6 days were considered: 11, 13 November and 1, 8, 14 and 20 December 2011 for a total of 14 precipitating events. For both campaigns the X-band polarimetric radar, manufactured by Gematronik (Germany) has been employed with the main characteristics: Magnetron with 35 Kw per channel, simultaneous horizontal and vertical polarization, pulse width of 0.5 μs, operative pulse repetition frequency (PRF) of 1500 Hz, 1.8 m antenna diameter, 1.3° beam width, operation frequency of 9.375 GHz, 150 m of range resolution and maximum distance of 100 km.

The Vale do Paraíba campaign occurred in São Paulo State in an elevated valley between the Serra da Mantiqueira and Serra do Mar mountain ranges. This field campaign had the longest duration for the CHUVA experiment, with an Intensive Observation Period (IOP) starting from 1 November to 22 December 2011, followed by a second period with less intensive measurements through 31 March 2012. The site strategy is indicated in Figure 1, where the X-band polarimetric radar was installed near São José dos Campos (Lat. 23°12′31.33″S, Lon. 45°57′7.87″W, 650 m ASL) above the roof of the UNIVAP building (Vale do Paraíba University) being approximately 82 km inland from the ocean. The radar scanning strategy produced a volume scan with 13 elevations (varying from 1 to 25 degrees) with a scan time of 6 min. Seven measurement sites (called UNIVAP, CTA, IEAV, Jambeiro, CESP, Pousada and Caragua) were established and equipped with rain gauges located at 9, 11, 22, 43, 51 and 75 km from the radar site, respectively, along a perpendicular line towards the ocean (Figure 1).

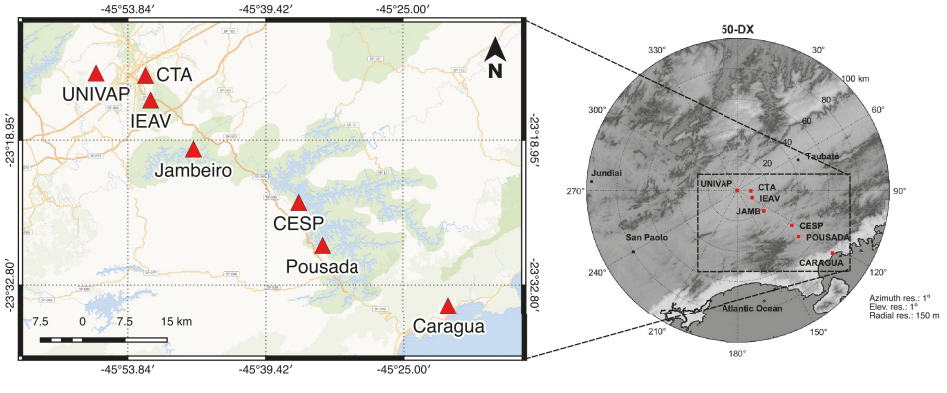


Figure 1. Position of the X-band radar and of the rain gauges indicated with a thumbtack in each site, during the Vale do Paraíba campaign.

The main rainfall systems that were observed during the campaign were caused by the penetration of cold fronts, local convection and organized mesoscale systems. The presence of a 500 hPa trough to the east (first half of the period) and west (second part) was responsible for the atmospheric moisture flow over the region and, consequently, the precipitation regime. Several thunderstorm events, some associated with hail, were reported during the campaign [10].

During the Manaus campaign, the X-band polarimetric radar was installed in the Amazon rainforest about 60 km from Manaus, between the Negro and Amazon rivers (Lat. 3°12′46.86″S, Lon. 60°35′53.92″W, 69 m ASL). The X-band radar scanning strategy produced one volume scan with 15 elevations (varying from 0.5 to 30 degrees) every 10 min. Two measurement sites (called T3 and Manacapuru) were established and equipped with rain gauges. T3 is located in the same position of the radar site while Manacapuru is approximately 10 km from the radar site (Figure 2). The field campaign occurred in two IOPs, the first happened from 13 February to 31 March 2014 during the wet season, and the second one between 1 and 30 September 2014 at the end of the dry season.

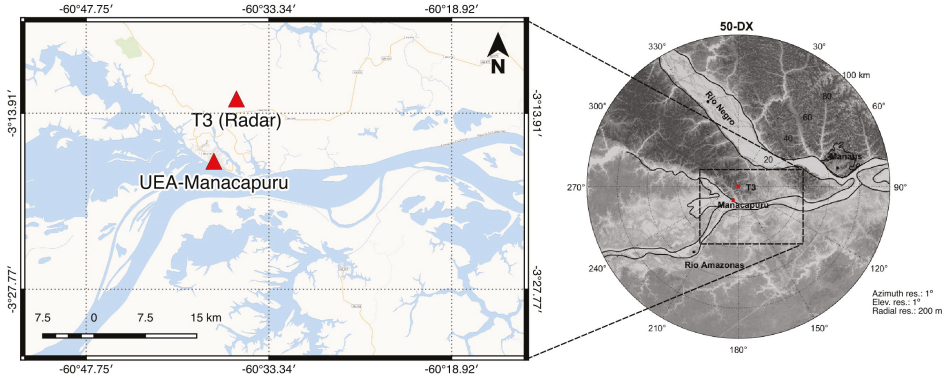


Figure 2. Position of the X-band radar and of the rain gauges indicated with a thumbtack in each site, during Manaus campaign.

The austral winter corresponds to the dry season in most of the Amazon region, although it represents a rainy maximum for the far northwest of the basin [11]. Well-defined wet and dry seasons are associated with the so-called South American Monsoon System (SAMS) [12].

2.2. Radar Data Quality Index

To perform a reliable validation procedure, the common validation methodology developed by the H-SAF Precipitation Product Validation Group (PPVG) was applied to the CHUVA radar data. The proposed scheme aims to compensate or, at least, minimize or eliminate those uncertainties. Moreover, a quality indicator for each source of error was introduced through appropriate tests. These quality matrices are composed by partial indexes that will be part of an overall data quality indicator as shown in Figure 3. In this figure, the main error sources taken into account are shown: ground clutter, partial beam blocking (PBB), range distance, non-uniform vertical profiles of reflectivity (VPR), differential phase processing and rain induced attenuation [8].

Any contamination of the radar signal by non-precipitation echoes, including returns from ground is considered in the $q_{clutter}$ index. In order to eliminate these effects, the raw volumetric data is treated trough a filtering technique based on a fuzzy logic approach. More details about this methodology can be accessed at [13]. The Figure 4 shows an example of clutter correction applied for an event occurred on 1 December 2011 in the Vale do Paraíba campaign.

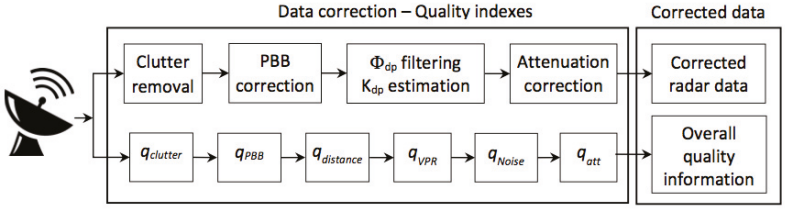
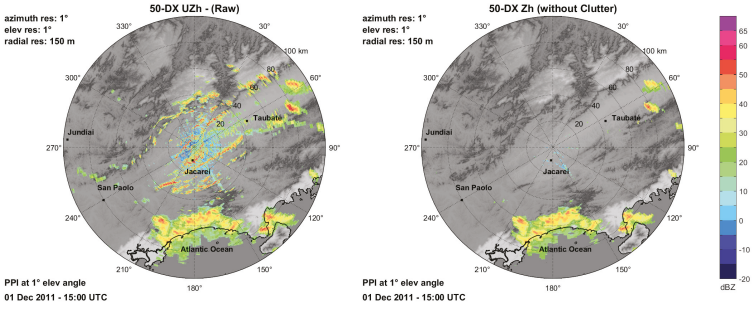


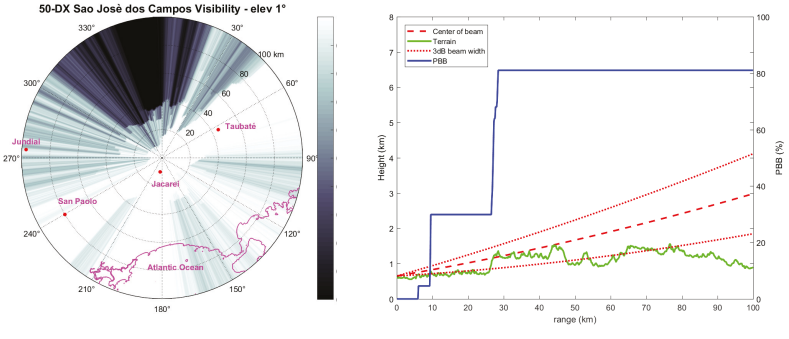
Figure 3. Weather radar data processing chain.



(**a**) (**b**)

Figure 4. PPI, at lowest elevation angle of (**a**) raw reflectivity and (**b**) reflectivity filtered to remove the clutter contamination on 1 December 2011 in the Vale do Paraíba campaign.

Another type of contamination is when the radar beam intercepts an obstruction. In this case, two situations are possible: (1) only part of the beam cross section illuminates the intercepted topography (partial blockage), or (2) the radar beam is completely blocked (total blockage). The beam blockage filtering approach q_{PBB} applied in this study is described in [14]. An example is shown in Figure 5 of the visibility map obtained at the lowest elevation angle.



(**a**) (**b**)

Figure 5. (**a**) Map of radar visibility in Vale do Paraíba (black, no visibility and white, full visibility) for the 1° elevation and (**b**) profile of the beam, PBB and DEM along a given azimuth (330 degrees). Maximum range is 100 km.

The radar data quality decreases with the beam broadening with the distance **r** from the radar. This effect cannot be corrected, however the data range-related deterioration can be determined quantitatively and taken into account in the $q_{distance}$. Following the approach proposed by [15] we adapted the equation from [6] considering $q = 0.5$ for $r \geq r_{max}$ to avoid an abrupt decrease in the quality index. The value 134 (km) is set so that for $r = 100$ (the radar range) we obtain $q = 0.5$ for $r \geq r_{max}$. It was also introduced a square root operation and the $q_{distance}$ index can be calculated using a non-linear function, as in following Equation (1):

$$q_{\text{distance}} = \begin{cases} 0.5 & \text{for } r \geq r_{max} \\ 1 & \text{for } r \leq r_{min} \\ \sqrt{\dfrac{134 - r}{134 - r_{min}}} & \text{for } r_{min} < r < r_{max} \end{cases} \quad (1)$$

where r_{max} can be set to 100 km and $r_{min} = \Delta r/2$ (Δr is the radar range resolution). The r_{max} is the radius of radar domain, the maximum range is 100 km. Range resolution is the ability of a radar system to distinguish between two or more targets on the same beam but at different ranges. It depends mainly on the width of the transmitted pulse and for this radar it is 150 m ($\Delta r = 150$ m). The square root is introduced in order to ensure that the quality does not drop too fast as the range distance increases.

Reflectivity values within the layer close to the 0° isotherm (referred to as melting layer or bright band) are overdetermined since it consists mainly of water-coated non-Rayleigh scatterers. Because of that, the most important parameter that defines the non-uniform vertical profiles of reflectivity (VPR) is the height of the freezing level (FL). The q_{VPR} is estimated following the Friedrich approach and taking into account the freezing level height, the layer between 200 m above the FL and 500 m below it, the beam width and the antenna elevation [15].

A polarimetric radar system provides measurements of the total differential phase Φ_{DP} that is the sum of the differential propagation Φ_{DP} and the backscatter phase δhv. Only the propagation component is considered for attenuation correction and for rainfall estimation purposes, with K_{DP} being related to the range derivative of Φ_{DP}, which is affected by system noise, offset, and potential aliasing problems. In order to handle this issue (q_{noise}), an Iterative moving-window range Finite Derivative scheme (IFD) approach is used and can be summarized through flow chart of Figure 6 and Equation (2). More information about this scheme can be accessed at [13].

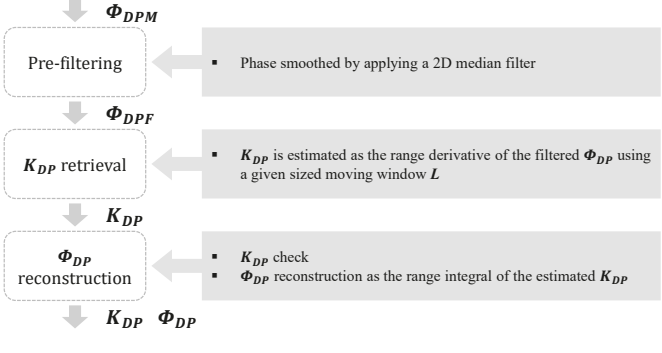


Figure 6. Weather radar data processing chain.

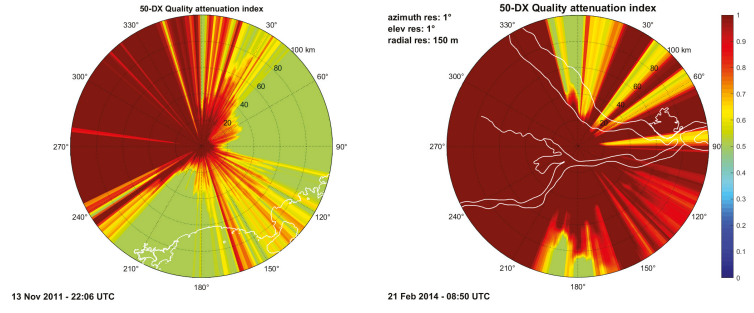
$$q_{\text{noise}} = \begin{cases} 1 & \text{if } K_{DP} \geq 0.5 \\ 2K_{DP} & \text{if } 0 < K_{DP} \leq 0.5 \\ 0 & \text{if } K_{DP} < 0 \end{cases} \quad (2)$$

Usually, there are a variety of possible solutions for attenuation correction in dual-polarization radars. All of them are based on the use of differential phase shift [16]. In the common polarimetric approach, for compensating rain path attenuation, the specific attenuation (AH) assumes a linear relationship with the specific differential phase (K_{DP}) [17]. This approach is applied for frequencies ranging from 5 GHz to at least 19 GHz, which includes both the C (5.6 GHz) and X-band (9.4 GHz) radars. The quality index q_{Att} associated with rain path attenuation can be defined as Equation (3):

$$q_{\text{att}} = \begin{cases} 1 & \text{for } PIA < PIA_{min} \\ 0.5 & \text{for } PIA > PIA_{max} \\ \dfrac{2PIA_{max} - PIA_{min} - PIA}{2(PIA_{max} - PIA_{min})} & \text{for } PIA_{min} \leq PIA \leq PIA_{max} \end{cases} \quad (3)$$

where PIA_{min} = 3 dB and PIA_{max} = 15 dB.

Finally, the overall radar data quality Q can be retrieved by combining all the considered quality indicators. All the partial quality matrices are used in a multiplicative combination. Figure 7 shows an example of the attenuation correction map in the Vale do Paraíba and Manaus campaigns. Additional examples of the overall quality index will be shown in Section 3.



(**a**) (**b**)

Figure 7. (**a**) Attenuation correction map at lowest elevation angle associated with radar in (**a**) Vale do Paraíba on 13 November 2011 at 22:06 UTC (center) and (**b**) Manaus on 21 February 2014 at 8:50 UTC (right).

2.3. Rainfall Estimation from Radar Data

Beyond the quality index procedure described in the previous section, another important aspect to be taken into account is that there are different algorithms in use for radar-based rainfall retrievals. Traditionally, the simplest rainfall relation is the Z-R relation, where rainfall (R) is estimated from reflectivity (Z). With the availability of polarimetric variables, it is possible to exploit several combinations (e.g., Z and Z_{DR}, among others) towards an optimal rainfall rate estimation. From the reflectivity (Z), the rainfall can be calculated by means of a power-law type relationship in Equation (4):

$$R(Z_H) = aZ_H^b \quad (4)$$

where R is rainfall rate in mm h^{-1} and Z_H in mm^6 m^{-3} units, while a and b are two coefficients whose value depends on the type of precipitation. The use of polarimetric quantities clearly provides better results in terms of precipitation estimation, being able to employ more than the reflectivity Z_H also the specific phase K_{DP} and the differential reflectivity Z_{DR} in various combinations. For the precipitation estimation, the *R*-*Z* and the *R*-K_{DP} relationships were adopted. Indeed, the algorithms that estimate

rainfall from K_{DP} are particularly attractive at wavelengths such as X-band because they are derived from phase measurements and they are unaffected by absolute calibration error and attenuation caused by precipitation along the propagation path [8]. The R-K_{DP} relationship can be written as Equation (5):

$$R(K_{DP}) = c \cdot (|K_{DP}|)^d \cdot sign(K_{DP}) \quad (5)$$

For the parameterization of $R(Z_H)$ and $R(K_{DP})$ (Equations (4) and (5)), the approach from [18] (MP48) and from CHUVA campaign in Fortaleza [19] (SC12) were used respectively. Another pair of coefficients is adopted for R-K_{DP} model, using [8] (BC01). Additionally, the parameters for $R(Z_H)$ and $R(K_{DP})$ are summarized in Table 1.

Table 1. Parameters of $R(Z_H)$ and $R(K_{DP})$ power law relations for the CHUVA X-band Radar.

	$R(Z_H)$			$R(K_{DP})$	
	a	b		c	d
MP48	0.0208	0.680	SC12	15.813	0.774
			BC01	19.193	0.850

The use of radar reflectivity for radar-based rainfall retrievals is frequently subject of underestimation [6], especially when the lowest beam map (LBM) is used. In order to reduce this effect, an additional polarimetric rainfall algorithm was proposed by Vulpiani [20]. The technique is based on the combination of reflectivity factor and specific differential phase and considers two types of weighting. The first consideration uses the K_{DP} gradually with increasing rainfall intensity, the combined algorithm takes the form of a weighted sum as Equation (6):

$$R_q = q_{\text{noise}} \cdot R_K + (1 - q_{\text{noise}}) \cdot R_Z \quad (6)$$

where R_Z and R_K are the rainfall estimates obtained by applying specific power laws (Equations (4) and (5)) to the lowest non-shielded radar bin of Z_H and K_{DP}, respectively. The coefficients that were adopted to derive R_K and R_Z are shown in Table 1, while the weight is the quality index q_{noise} (Equation (2)). The second consideration uses a combined polarimetric rainfall algorithm with another type of weighting between the contributions of Z_H and K_{DP}, as Equation (7):

$$R_q = \frac{q_{\text{loss}} \cdot R_Z + q_{\text{noise}} \cdot R_K}{q_{\text{loss}} + q_{\text{noise}}} \quad (7)$$

where R_K and R_Z parameters are the same of Equation (6), while the weights are the quality indexes q_{loss} ($q_{loss} = q_{pbbx} \times q_{att}$) and q_{noise} indicated in Equation (2). Therefore, the different algorithms used with the various coefficient combinations are summarized as follow:

1. R_Z : (4) with MP48 coefficients.
2. R_{K1}: (5) with SC12 coefficients.
3. R_{K2}: (5)with BC01 coefficients.
4. R_{q1}: (7) with MP48 and SC12 coefficients.
5. R_{q2}: (7) with MP48 and BC01 coefficients.
6. R_{q1Vu15}: (6) with MP48 and SC12 coefficients.
7. R_{q2Vu15}: (6) with MP48 and BC01 coefficients.

The results related with the different algorithms will be shown in Section 3.

2.4. Satellite Products and Dataset Generation

The passive microwave (PMW) precipitation products within the EUMETSAT H-SAF ([21]) are based on the development and refinement of retrieval techniques exploiting all available radiometers

in the GPM constellation. In this context, operational PMW precipitation products for the different radiometers are being released within H-SAF. They are based on two approaches ([22]): the physically based Bayesian Cloud Dynamics and Radiation Database (CDRD) algorithm ([23,24]) for conically scanning radiometers and the Passive microwave Neural network Precipitation Retrieval algorithm (PNPR) for cross-track scanning radiometers ([25,26]). Three PMW H-SAF products were considered in this study: H01 (CDRD approach applied to SSMIS), H02 (PNPR developed for AMSU/MHS), and H18 (PNPR adapted to ATMS).

The algorithms are based on the use of a cloud-radiation database made up of thousands of microphysical-meteorological profiles derived from cloud-resolving model simulations of different precipitation events including 60 simulations over the European/Mediterranean area ([23]) and 34 simulations over Africa and Southern Atlantic. The main features of each product can be accessed in detail in the respective references cited above.

To perform the validation over Brazil, we had to acquire all the input data (brightness temperatures in TB) for the the CHUVA experiments timeframe, and then, to process the H-SAF products to extend the coverage of the retrievals in order to include the whole country (extended to 75°N–60°S and 80°W–80°E). The inputs for the H01 were the SSMI/S orbits of the DMSP F16, F17 and F18 satellites. For the H02 algorithm, the input files were the AMSU-A, AMSU-B and MHS orbits from NOAA-18, NOAA-19, MetOp-A and MetOp-B. In addition, lastly, we used the ATMS orbits from the Suomi-NPP satellite as input for the H18 algorithm. Though we have analyzed the three products: H01, H02 and H18 for the Manaus campaign, we only analyzed the H01 and H02 products for the Vale do Paraíba campaign because the ATMS data was not available in 2011.

2.5. Application of the Common Validation Code (CVC)

The common validation code (CVC) developed by the Precipitation Product Validation Group (PPVG) [7] enables implementation of a common validation procedure to make the validation results comparable. The products to be validated differ in terms of retrieval technique, spatial and temporal resolutions. Therefore, each product requires a specific validation procedure. The methodology can be divided on the following general steps:

- Ground data error analysis;
- Upscaling of radar data to match the satellite product nominal resolution;
- Temporal matching of precipitation products (satellite and ground);
- Application of evaluation statistical methods (continuous and multi-categorical) to all available overpasses, for each pixel pair (satellite-ground).

In order to use the CVC on the Brazilian radar data, the code had to be adapted by including the new radar coordinates (Vale do Paraíba and Manaus) and changing the reading routine to the new radar data. The CVC was configured to match a maximum temporal difference between satellite and radar of 16 min. The radar data was upscaled to the satellite product nominal resolution, considering the antenna pattern (Gaussian function), viewing geometry, and scanning strategy (conical and cross-track) of the MW radiometers. As the radar data was filtered beforehand (quality control), the pixels with low quality were eliminated in the upscale processing. To investigate the performance of the precipitation products, the statistical scores commonly used in the pixel-based validation by the H-SAF PPVG were considered.

3. Results

In this section, we present the results of the evaluation of rain gauge measurements and the performance radar-based rainfall algorithms for one case study of Vale do Paraíba and Manaus, respectively. Sequentially, we performed the validation of the H01, H02 and H18 algorithms, breaking it down into two phases, the statistical evaluation and pixel-by-pixel analysis. In order to summarize

the results for all case studies, we chose to exhibit one representative event for each campaign; however, the results of the remaining cases will also be discussed further ahead.

3.1. Vale do Paraíba Campaign: 13 November 2011

This precipitating event refers to the occurrence of extensive convection followed by stratiform rainfall. Figure 8 shows the 24-h accumulation (from 00:00 to 24:00 UTC) derived from the distinct rainfall algorithms considered in this study. Each map corresponds to a specific rainfall technique and includes the cumulative values recorded by rain gauges on the left of each map. In comparison to rain gauges, the radar QPE based on R_Z estimator (Figure 8a), shows clear signs of underestimation, especially regarding the convective cores. On the other hand, the K_{DP}-based algorithms have a similar pattern to those recorded by rain gauges.

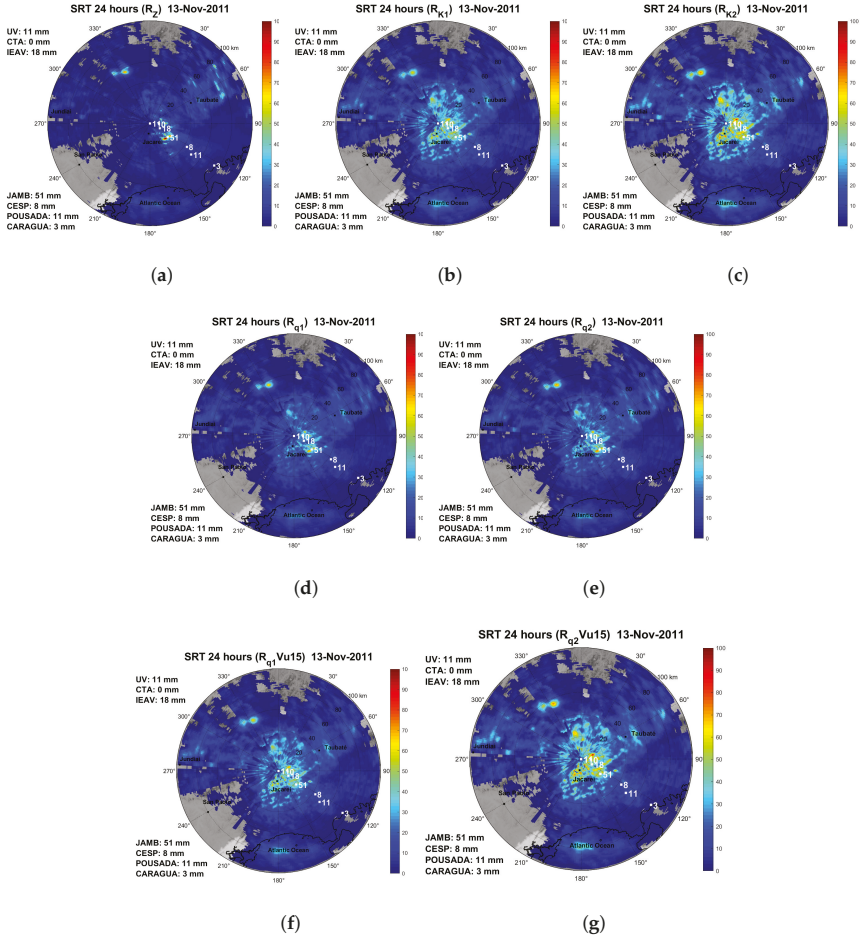


Figure 8. Surface Rainfall Total (SRT) 24 hourly accumulated using (**a**) R_Z model, (**b**) R_{K1} model, (**c**) R_{K2} model, (**d**) R_{q1} model, (**e**) R_{q2} model, (**f**) R_{q1Vu15} model and (**g**) R_{q2Vu15} model for the event which occurred on 13 November 2011. The corresponding values accumulated by the rain gauges are shown in the map.

3.2. Manaus Campaign: 8 March 2014

This event was characterized by extensive and intense stratiform precipitation in Manaus. Figure 9 presents the different algorithms with the rain gauge records indicated at the bottom-left of each map. Similarly of the Vale do Paraíba event the algorithm just based on reflectivity (R_Z) also presented a general underestimation of 24-h accumulation (Figure 9a) and the R_{q2Vu15} presented the best performance (Figure 9g).

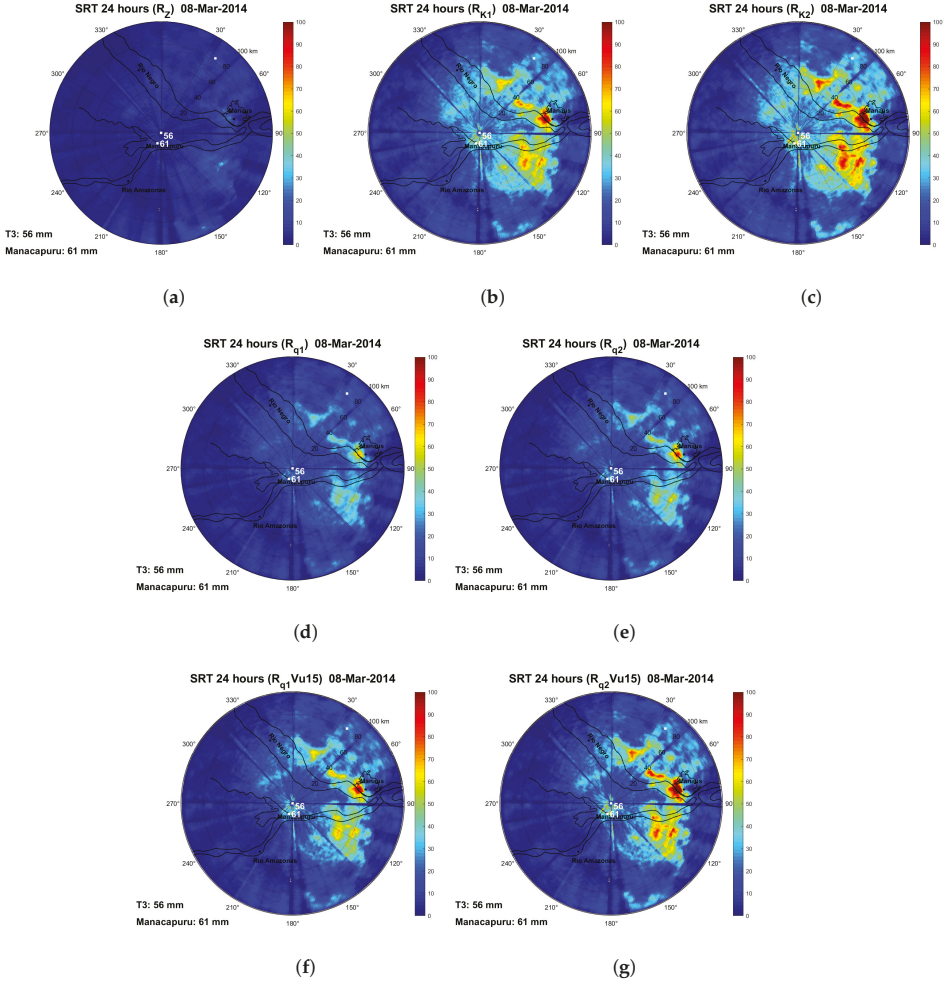


Figure 9. Surface Rainfall Total (SRT) 24 hourly accumulated using (**a**) R_Z model, (**b**) R_{K1} model, (**c**) R_{K2} model, (**d**) R_{q1} model, (**e**) R_{q2} model, (**f**) R_{q1Vu15} model and (**g**) R_{q2Vu15} model for the event which occurred on 8 March 2014. The corresponding values accumulated by the rain gauges are shown in the map.

A performance verification of the radar-based rainfall retrievals was made regarding the rain gauges (figures and tables were not shown here). The comparison considering the nearest value shows better results. In general, the radar-based rainfall R_Z (just based on reflectivity) presented

underestimation, the K_{DP}-based algorithms had good performance when compared with rain gauges measurements. The best estimator was the R_{q2Vu15}, because it presented the perfect correlation coefficient score, which is one. And bias, mean absolute error and root mean square not higher than 0.11, 0.13 and 0.41, respectively.

3.3. Validation of H-SAF Precipitation Products

The second part, which is the main focus of this study, is to identify the performance of the H-SAF products in order to provide the algorithm developers with information on the limitations and issues of the retrievals over specific regions in Brazil. The case studies under analysis are the same 14 cases (for both campaigns) that were considered previously on the radar data quality analysis. As mentioned before, three satellite rainfall products were analyzed: H01 (CDRD approach applied to SSMIS), H02b (PNPR v1 developed for AMSU/MHS), and H18 (PNPR v2 adapted to ATMS). We present a statistical analysis based on continuous and dichotomous statistical scores computed within the CVC, along with an in-depth analysis of the selected cases, followed by a pixel by pixel analysis. The considerations concerning the acquisition of the matching pairs (radar $\times$ satellite) are exposed in Section 2.5.

3.3.1. Statistical Evaluation

Because of its best estimation performance, the R_{q2Vu15} algorithm was chosen for the statistical evaluation. In order to investigate the impact of different quality indexes in the radar retrievals, we performed a sensitivity analysis (not shown here). According to this analysis, we found that a quality index value equal or greater than 0.7 could be considered a good compromise between the desired performance of the statistical scores and the size of the dataset (that guarantees the reliability of the results), i.e., the number of pixels in the sample. The statistical scores used in the validation procedure are presented in Tables 2 and 3. The continuous statistical scores were computed for the pixels in which both radar and satellite give rainfall estimates larger than 0.25 mm/h (hits only). On the other hand, all pixels are considered for the multicategory scores. The number of pixels for each algorithm is different because it depends on the number of available satellite overpasses.

For Manaus, the events under analysis counted a total of 33 overpass matchings (which means correspondence in time and space for both the satellite and radar measurements) for H01, 49 matchings for H02 and 13 matchings for H18. Since ATMS is aboard a single satellite, the number of overpasses over the region of interest is less than the other sensors (consequently, lower number of matched pixel pairs).

Table 2. Continuous statistical scores.

Score	Perfect Score	Calculation
Mean error or bias (ME)	0	$ME = \frac{1}{N}\sum_{k=1}^{n}(sat_k - obs_k)$
Standard deviation (SD)	0	$SD = \sqrt{\frac{1}{N}\sum_{k=1}^{n}(sat_k - obs_k - ME)^2}$
Root mean square error (RMSE)	0	$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{n}(sat_k - obs_k)^2}$
Fractional standard error (FSE)	0	$FSE = \frac{\sqrt{\frac{1}{N}\sum_{k=1}^{n}(sat_k - obs_k)^2}}{\frac{1}{N}\sum_{1}^{n} obs_k} = \frac{RMSE}{\frac{1}{N}\sum_{1}^{n} obs_k}$
Correlation coefficient (CC)	1	$CC = \frac{\sum_{k=1}^{n}(sat - \bar{sat})(obs - \bar{obs}))}{\sqrt{\sum_{k=1}^{n}(sat - \bar{sat})^2}\sqrt{\sum_{k=1}^{n}(obs - \bar{obs})^2}}$

Table 3. Multicategory scores.

Score	Perfect Score	Calculation
Probability of Detection (POD)	1	$POD = \frac{hits}{hits + misses} = \frac{hits}{observed\ yes}$
False alarm rate (FAR)	0	$FAR = \frac{false\ alarms}{hits + false\ alarms} = \frac{false\ alarms}{forecast\ yes}$
Critical success index (CSI)	1	$CSI = \frac{hits}{hits + misses + false\ alarms}$

Table 4 presents continuous scores for the algorithms in analysis, the number inside the brackets refers to the number of matched pixel pairs for each algorithm. All algorithms tend to overestimate the X-band radar estimates. The H01 presented the largest values for ME, RMSE and FSE. In general H02 had slightly better scores than H01, such as lower mean rainfall rate values, ME, RMSE, SD and FSE. Finally, the H18 presents intermediate scores in relation to the other algorithms, having just the smallest FSE. Regarding multicategory scores (Table 5), it is worth noticing that both algorithms based on the neural network approach presented better detection skills than H01, the highest POD (0.96 and 0.81), the lowest FAR (0.47 and 0.39) and highest critical CSI (0.51 and 0.53), respectively.

Table 4. Statistical continuous scores for Manaus for H01, H02 and H18. The number inside parenthesis represents the number of matched pixel pairs for each algorithm.

Algorithm	Sat. Mean	Rad. Mean	ME	RMSE	SD	FSE	CORR
H01 (406)	4.60	0.93	3.67	5.20	3.69	5.59	0.40
H02 (631)	3.78	1.18	2.59	5.11	4.40	4.32	0.32
H18 (140)	4.53	1.46	3.07	5.10	4.07	3.48	0.34

Table 5. Multicategory scores for Manaus for H01, H02 and H18. The number inside parenthesis concerns the number of matched pixel pairs for each algorithm.

Algorithm	POD	FAR	CSI
H01 (3149)	0.75	0.54	0.39
H02 (1889)	0.96	0.47	0.51
H18 (916)	0.81	0.39	0.53

The Vale do Paraíba campaign counted a total of 27 overpass matchings for H01 and 31 matchings for H02. The respective statistical scores are presented in Tables 6 and 7. We can see from Table 6 that H01 and H02 algorithms presented similar mean rain rates, slightly lower than the radar estimates, thus determining small and negative values for ME. These results can be attributed to either the better performance of the satellite products for this region (at least in terms of bias), or the better quality of the ground-radar estimates over this region. Compared to Manaus region, the multicategory scores for Vale do Paraíba presented lower FAR values (0.39 for H01 and 0.21 for H02), lower POD values (0.49 for H01 and 0.43 for H02) and similar CSI values

Table 6. Continuous scores for Vale do Paraíba for H01 and H02. The number inside parenthesis concerns the number of matched pixel pairs for each algorithm.

Algorithm	Sat. Mean	Rad. Mean	ME	RMSE	SD	FSE	CORR
H01 (181)	2.53	2.89	−0.35	4.00	3.98	1.38	0.44
H02 (124)	2.45	2.76	−0.30	3.56	3.55	1.29	0.37

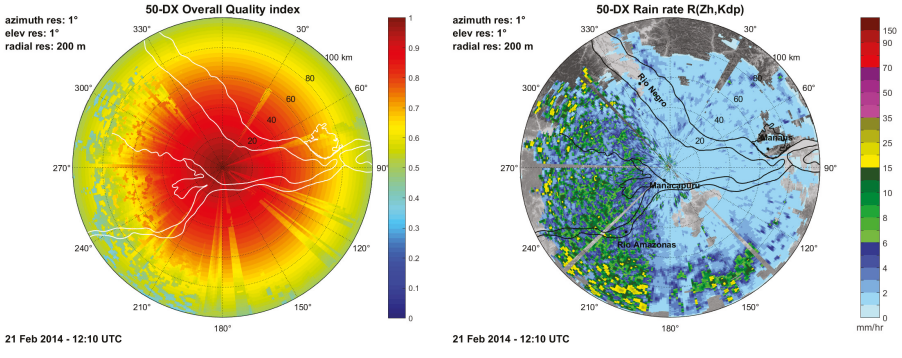
Table 7. Multi-category scores for Vale do Paraíba for H01 and H02. The number inside parenthesis concerns the number of matched pixel pairs for each algorithm.

Algorithm	POD	FAR	CSI
H01 (2420)	0.49	0.49	0.33
H02 (991)	0.43	0.21	0.39

3.3.2. Pixel by Pixel Analysis

In order to further analyze the algorithms performance, a case study analysis for each product and each campaign, with pixel by pixel comparison, was carried out.

A case study for Manaus, which occurred on 21 December 2014 was presented, with the overall quality index shown in Figure 10a and the radar rainfall field in its original resolution in Figure 10b. On this day, a well defined squall line approached the X-band radar region. The systems persisted on the region from the morning around 08:20 UTC to the afternoon 15:40 UTC. It is possible to see the presence of convective cores with rainfall rate upwards to 35 mm/h in the regions between the north-west and the south-west quadrants in Figure 10b. The remaining regions are dominated by light and stratiform precipitation with rainfall rates going up to 6 mm/h. In Figure 11a it is possible to see the filtered radar data by applying the quality index threshold at 0.7 and upscaled to the satellite native grid, and in Figure 11b the respective H01 rainfall retrieval is shown. The light rain rate provided by the radar (light blue), is associated with moderate precipitation values by H01 (shades of green). Additionally, the moderate rain rates from radar (green and dark blue) is strongly overestimated by the H01 (orange to red) algorithm. We can say that H01 has a general tendency to overestimate all rainfall classes.



(**a**) (**b**)

Figure 10. (**a**) Overall radar quality index and (**b**) Rain rate from radar on 21 February 2014 at 12:10 UTC.

For H02, one overpass for the case which occurred on 8 March 2014 (at 05:30 UTC) is shown. Convective cells are visible along the border of the Rio Negro River, followed by smaller convective cells over the Southeast quadrant (Figure 12). Figure 13a,b show that the H02 algorithm provides precipitations from light to moderate values (shades of green) in pixels where the radar detects absence of precipitation (gray pixels). This tendency to produce a large area of precipitating pixels is related to the precipitation screening, which is the algorithm module where potential precipitating pixels are selected. This feature is related to the relatively high FAR scores (shown in Table 7).

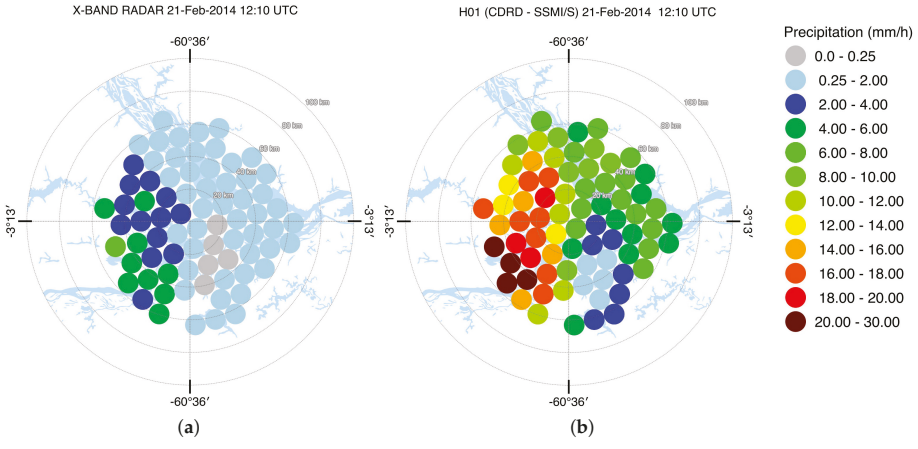


Figure 11. (**a**) Rain rate from radar upscaled to the satellite grid and (**b**) H01 rain rate retrieval in Manaus on 21 February 2014 at 12:10 UTC.

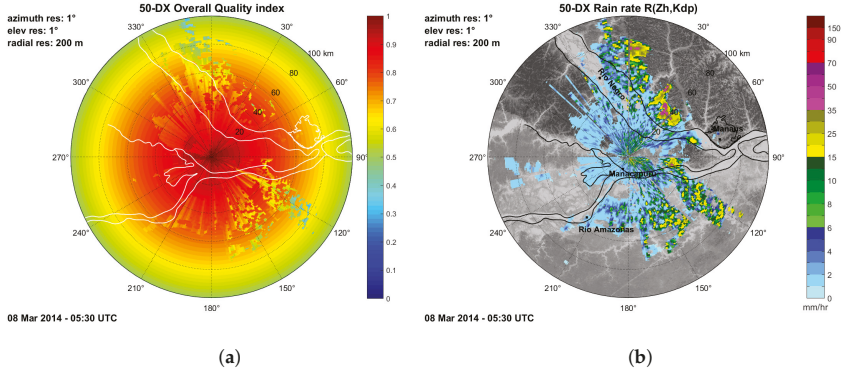


Figure 12. (**a**) Overall radar quality index and (**b**) Rain rate from radar on 8 March 2014 at 05:30 UTC.

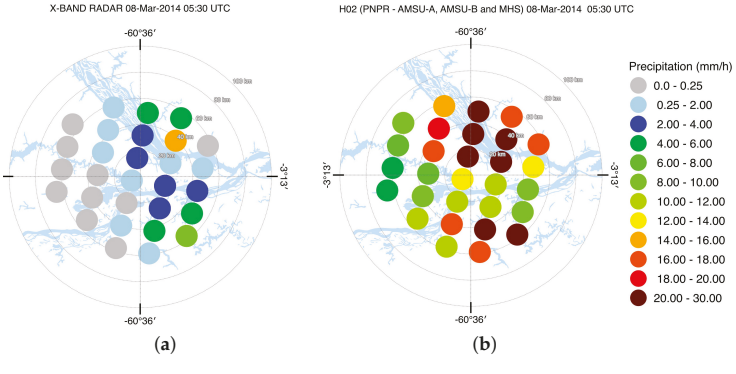
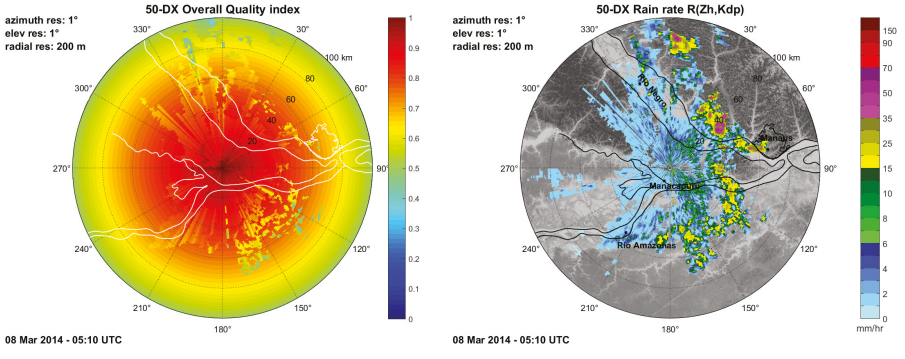


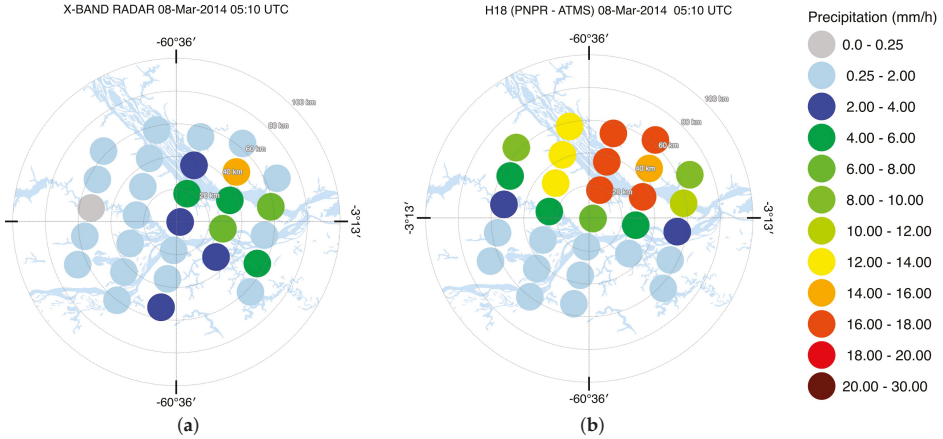
Figure 13. (**a**) Rain rate from radar upscaled to the satellite grid and (**b**) H02 rain rate map in Manaus on 8 March 2014 at 05:30 UTC.

The example for H18 shown in Figures 14 and 15 is for the same event presented in H02, on 8 March 2014 but at 05:10 UTC. Despite the fact that the satellite overpasses (H02 and H18) were very close in time (20 min of difference) it was enough for modifications in the precipitation field to appear. The H18 precipitation pattern is similar to H02 because they are based on the same precipitation screening method. It is worth pointing out that the screening of precipitation is equal for all products in analysis in this study and it is applied over all background surfaces, except over desert. It is very likely that the screening procedure is not well tuned for the atmospheric conditions of the Amazon region, characterized by the high water vapor content. On the other hand, in the area with the most intense convective cores, the H18 tends to produce a lower overestimation rate when compared to H02. This aspect is related to the difference in the two retrieval algorithms (PNPR for AMSU/MHS and PNPR for ATMS) (as pointed out by [26], where H18 is based on just one neural network for all surface types, trained with a unique database with the additional channels in the water vapor absorption band at $TB_{183\pm3}$ ([26]).



(**a**) (**b**)

Figure 14. (**a**) Overall radar quality index and (**b**) Rain rate from radar on 8 March 2014 at 05:10 UTC.

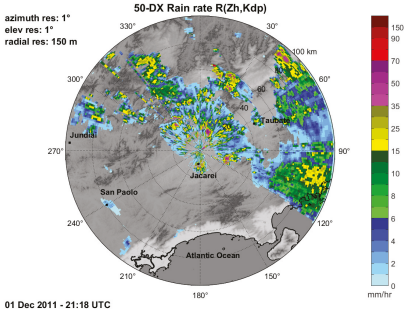


(**a**) (**b**)

Figure 15. (**a**) Rain rate from radar upscaled to the satellite grid and (**b**) H18 rain rate map in Manaus on 8 March 2014 at 05:10 UTC.

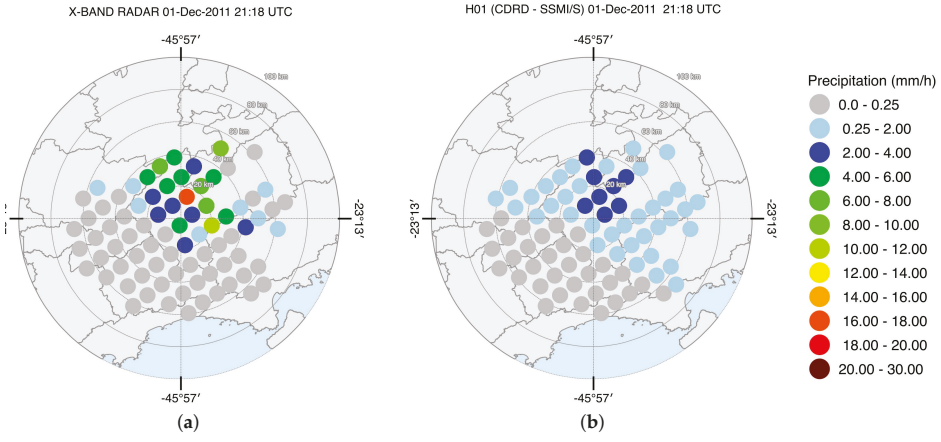
It is noticeable that the three algorithms have a tendency to overestimate the larger rain rates. The study in reference [27] evaluated the ability of two GPM rainfall algorithms (GPROF2014 and IMERG) in reproducing the main characteristics and the diurnal cycle of precipitation as observed by the S-band SIPAM radar, in the Manaus region during the CHUVA campaign. The authors have found similar results where GPROF2014 (GMI) presents large overestimation of the rain rate volume and occurrence greater than 10 mm h^{-1} during IOP1.

The first example for the Vale do Paraíba campaign refers to the presence of convective clusters associated with the SACZ (South Atlantic Convergence Zone). The coincident overpass between the satellite and the radar occurred at 21:18 UTC and shows the presence of a few intense convective cells along the Rio Negro River (Figure 16b). The overall quality index (Figure 16a) in this region is most affected by the blocking effects caused by two mountain ranges, the Serra da Mantiqueira (western) and Serra do Mar (eastern). Analyzing the upscaled maps (Figure 17) it is clear that, differently from Manaus, the precipitation pattern is well detected by the H01. Concerning the rainfall estimation, in this case, the tendency of the algorithm is to underestimate the highest precipitation rates seen by the radar. Despite this feature, the estimates are in better agreement with the radar than the estimates in the Manaus campaign (as shown by the lower ME for this region in Table 6).



(**a**) (**b**)

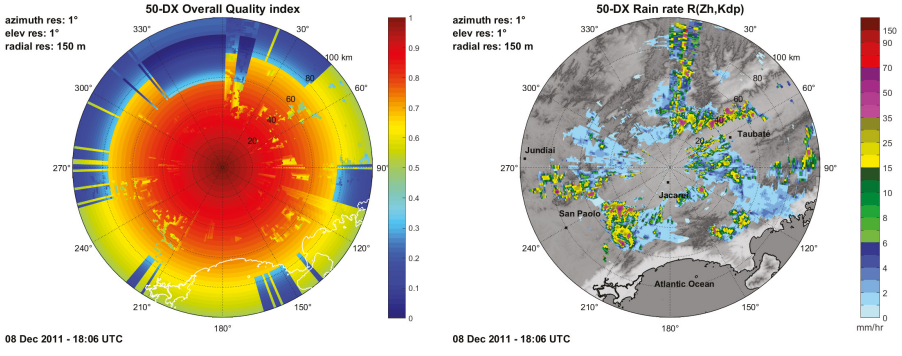
Figure 16. (**a**) Overall radar quality index and (**b**) Rain rate from radar on 1 December 2011 at 21:18 UTC.



(**a**) (**b**)

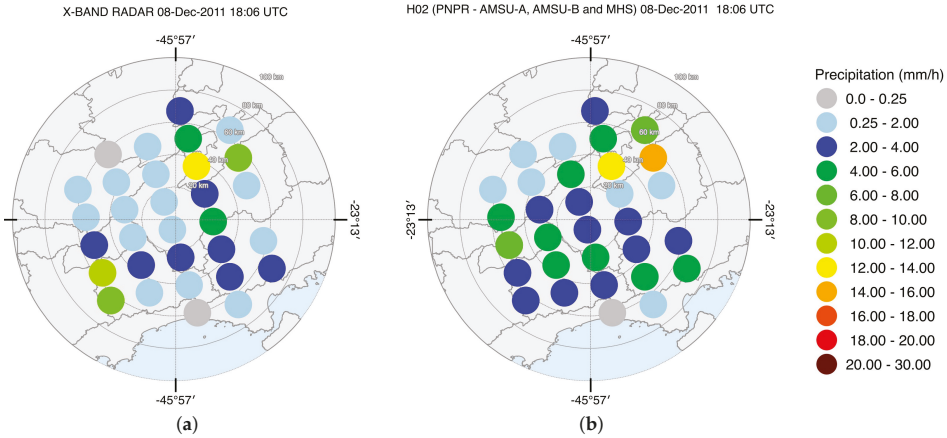
Figure 17. (**a**) Rain rate from radar upscaled to the satellite grid and (**b**) H01 rain rate map in Vale do Paraíba on 1 December 2011 at 21:18 UTC.

The second example for the Vale do Paraíba campaign (Figures 18 and 19) consists of the occurrence of local convection with sparse intense convective cells occurring predominantly in the afternoon (Figure 18b). During this event, the precipitation pattern is also quite close to the ground-based reference, which is also reflected on the lowest FAR values (Table 7). Concerning the rainfall estimation, in this case, the algorithm tends to underestimate the highest precipitation rates observed by the radar. However, the PMW estimates are in better agreement with the radar observations than the estimates generated in the Manaus campaign (as indicated by the lower ME for this region in Table 6).



(**a**) (**b**)

Figure 18. (**a**) Overall radar quality index and (**b**) Rain rate from radar on 8 December 2011 at 18:06 UTC.



(**a**) (**b**)

Figure 19. (**a**) Rain rate from radar upscaled to the satellite grid and (**b**) H02 rain rate map in Vale do Paraíba on 8 December 2011 at 18:06 UTC.

4. Discussion

We will discuss the main features of the H01 algorithm that mostly influence the results in Brazilian areas: the precipitation screening process, the surface classification and the representativeness of the cloud model simulations in the *a priori* database.

The first aspect, the precipitation screening process, selects the potentially precipitating pixels before the retrieval process. The screening used in the algorithms is based on the method described by [28], which uses the comparison of the TBs, which are the water vapor absorption band at $TB_{183\pm3}$ and $TB_{183\pm7}$ GHz and the TB_{53}GHz channel in the oxygen absorption band around 50 GHz. A detailed description of the implementation of the H-SAF products screening process can be accessed in [21,25].

Figure 20 present the radiative signatures from the $TB_{183\pm3}$, $TB_{183\pm7}$ and TB_{53}GHz channels for Manaus for the case study on 21 February 2014. Analyzing the radar area, we can notice that in Amazon region (Figure 20a,b) both water vapor absorption channels present a strong TB decay (minimal values around 160–180 K). This pattern seems to be associated with a change in the weight function peak of these channels in the presence of high water content, typical of the Amazon rainforest. As a consequence, an extensive area of precipitation can be seen in the screening map and in the final rainfall rates (Figure 20d,e). This can affect the efficiency of the screening in identifying potential precipitating pixels, especially when the clouds are not very deep as they appear to be during the wet season in the Amazon region.

On the other hand, for the Vale do Paraíba region (Figure 21), the channels 183 $\pm$ 7 and 183 $\pm$ 3 GHz do not present strong decay in TB inside the radar area and the screening map (Figure 21d) follows the same pattern that was observed in 183 $\pm$ 7 GHz channel ((Figure 21b). Differently from Manaus, we can see that the retrieval does not generate precipitation (gray pixels in Figure 21e) in the entire area provided by the screening as potentially precipitating (indicated by red in Figure 21d).

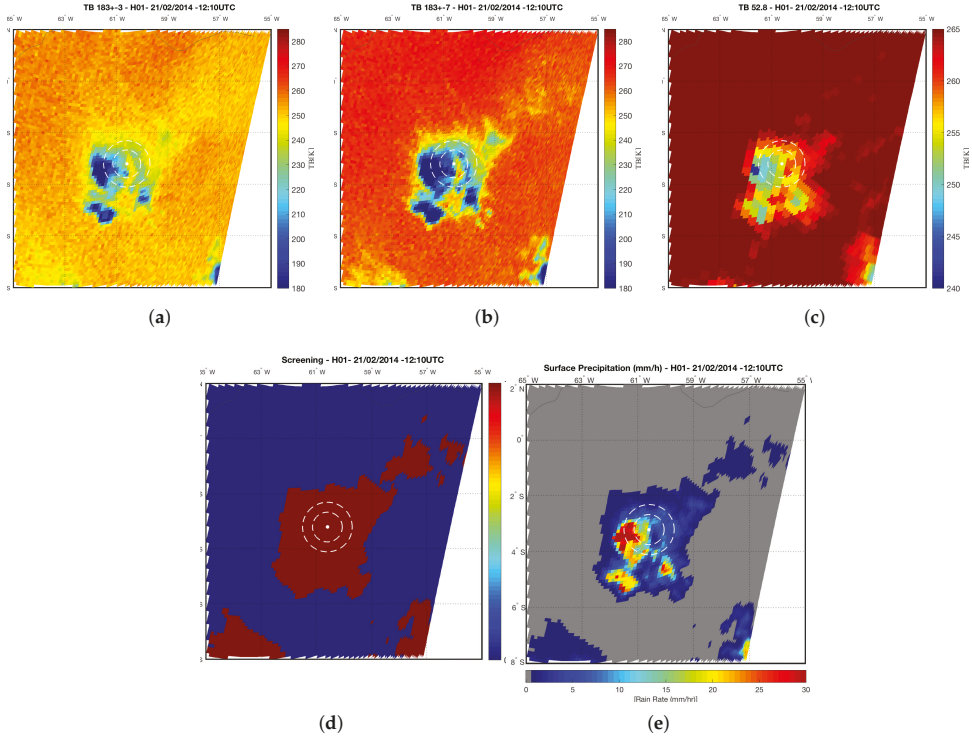


Figure 20. (**a**) $TB_{183\pm3}$ GHz, (**b**) $TB_{183\pm7}$ GHz, (**c**) TB_{53} GHz, (**d**) Screening of precipitation (0—no rain, 1—rain) and (**e**) Surface Precipitation (mm/h) on 21 February 2014 at 12:10 UTC in the Manaus region.

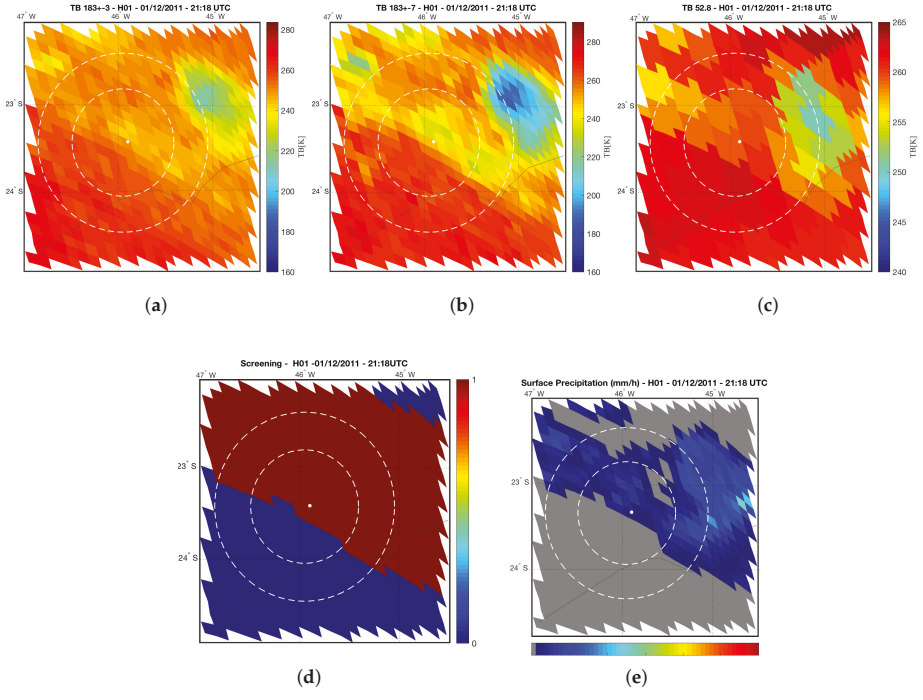
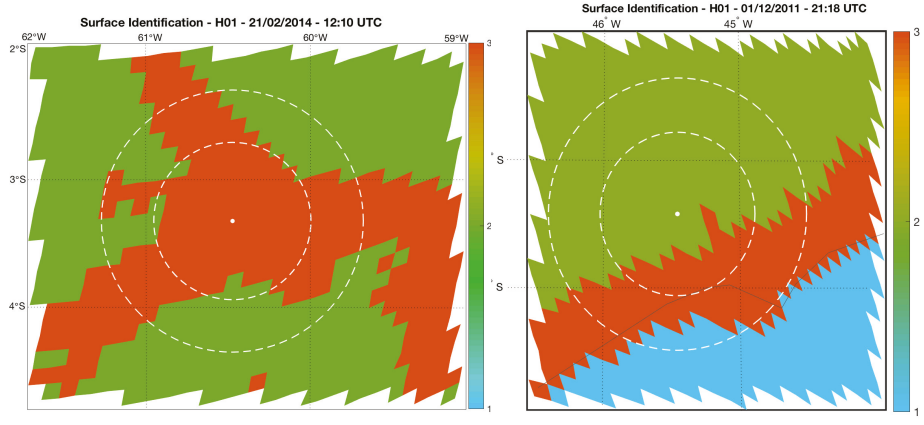


Figure 21. (**a**) $TB_{183\pm3}$ GHz, (**b**) $TB_{183\pm7}$ GHz, (**c**) TB_{53} GHz, (**d**) Screening of precipitation (0—no rain, 1—rain) and (**e**) Surface Precipitation (mm/h) on 1 December 2011 at 21:18 UTC in the Vale do Paraíba region.

In the Amazon region we have noticed that the surface characteristics were not properly classified. This region is dominated by large inland water bodies (Amazonas and Rio Negro rivers) and surrounded by vegetated land. Figure 22a shows that, in the radar area, the algorithm classifies the surface as coastal pixels, which is the least populated database, therefore, the least representative and usually affected by larger uncertainties. In the Vale do Paraíba region (Figure 22b) the algorithm considers large part of the radar area as vegetated land and the Bayesian approach look for land profiles, which leads to the choice of more appropriate rain rate profiles.

In the Bayesian approach, the surface classification impacts the choice of channels used in retrieval processing, which influences the selection of hydrometeor profiles in the *a priori* database and, ultimately, the rainfall rate retrieval.

Finally, the representativeness of the cloud model simulations in the *a priori* database, which is currently optimized for Europe and Africa, does not properly represent the typical microphysical and rainfall profiles of brazilian regions. Improving the database including simulations that represent the brazilian rainfall regimes could solve the issue.



(**a**) (**b**)

Figure 22. Surface Identification (1—Ocean, 2—Land, 3—Coast) for (**a**) Manaus and (**b**) Vale do Paraíba.

5. Conclusions

This work analyzed the data collected by mobile X-band polarimetric radars during two campaigns of the CHUVA project, both in the Vale do Paraíba and Manaus regions. First, in order to ensure that the verification of the H-SAF precipitation products is based on well-treated data, the ground radar data was submitted to a quality control procedure and the rainfall estimates were tested with different algorithms. The current polarimetric processing chain attempted to correct the error contributions of the different sources of uncertainty and provide the estimation of the reflectivity (Z_H) and the specific differential phase (K_{DP}), which contain the microphysical information required to perform attenuation correction and quantify the precipitation rate. Different polarimetric QPE algorithms have been coded and evaluated at a hourly time-step using independent rain gauges. The algorithms that were tested are: simple Z-R relationship (Marshall-Palmer, 1948 [18]) indicated as R_Z, algorithms based on K_{DP} solely (R_{K1} and R_{K2}), algorithms based on combination of K_{DP} and Z_H, whose weights are the quality indexes q_{loss} and q_{noise} (R_{q1} and R_{q2}), and algorithms based on combination of K_{DP} at medium-high K_{DP} (above 0.5° km^{-1}) and attenuation-corrected Z_H at low K_{DP} (below 0.5° km^{-1}) (R_{q1Vu15} and R_{q2Vu15}). The results confirmed the benefits brought by polarimetry to quantify radar rainfall retrievals and seems to indicate (considering the problems of a possible radome attenuation, which are not considered here), that the algorithm which best estimates the precipitation intensity was R_{q2Vu15}. Even though R_{K2} also provided good results, R_{q2Vu15} had better performance for both low and for high precipitation rates.

Regarding the results from the satellite algorithms validation, for the Manaus region, the CDRD algorithm (H01) tends to overestimate all rain rates classes (light to heavy). The PNPR algorithm for AMSU-A/AMSU-B/MHS sensors (H02) presents better POD than H01, but it also presents high FAR values. The PNPR for the ATMS sensor (H18) presents lower overestimation of heavy rain rates when compared to H02, probably due to the different neural network used on H18. It is worth considering that the sample size for this verification study (14 cases) was quite small, and a more extensive validation with a larger ground-based dataset would be suggested to perform a more comprehensive quality assessment.

All analyzed rainfall retrieval algorithms for the Amazon region showed high FAR values and larger precipitation patterns which are deeply related to the precipitation screening scheme. The screening seems to be substantially affected by the high water vapor content in this region. The H01 (CDRD for SSMIS—Bayesian approach), in specific, it was affected by highly variable surface

emissivities, impacting the selection of hydrometeor profiles in the *a priori* database and, ultimately, the rainfall rate retrieval. Moreover, we must take into account the uncertainties on precipitation estimates, which in the Bayesian approach are represented by the coefficients of error covariance matrices (for TBs, and ancillary and meteorological parameters) as analyzed in [24]) and in [29]. For the Vale do Paraíba region, both algorithms, H01 and H02, produced ME values that were quite close to zero (or negative) and lower FAR values (from 0.21 to 0.49) than Manaus. Unlike the Manaus case, the precipitation patterns were well detected and the estimations were in good agreement with the reference as indicated by the low ME values.

As for future plans, we intend to perform new verifications in a denser and more comprehensive sample for other Brazilian regions. Since we identified limitations in the precipitation screening scheme over the Amazon region, we plan on developing a new scheme that is able to handle the water vapour content in this region. That would require, furthermore, work on the development of a new surface classification capable of characterizing the surface diversity properly. The next steps must be made using the pluviometric precipitation dataset provided by the National Institute for Space Research (CPTEC/INPE) and also the dataset from new radars acquired by the National Center for Natural Disaster Monitoring and Alerts (CEMADEN). Lastly, for future work we intend to extend the cloud-radiation database used as *a priori* information in the algorithms in order to obtain radiative and microphysical profiles representative of brazilian rainfall regimes.

Author Contributions: L.M.C.d.A. and S.B. designed the study, conducted the analysis and wrote the manuscript. T.B., M.P., P.S. and A.C.M. provided support for data processing. The authors D.V., S.P., G.V., G.P., M.G. and S.D. contributed deeply to discussions, corrections and revisions, providing important feedback and suggestions.

Funding: The first author acknowledges the financial support of the National Council for Scientific and Technological Development (CNPq)-Brazil (Reference number-140715/2017-7 and Universal Call number 403987/2016-4) and the Coordination for the Improvement of Higher Education Personnel (CAPES) Brazil during her PhD studies and also the Science Without Borders (CSF) Program by the Sandwich Doctorate (SWE)(201237/2015-6) for the internship opportunity. The first and second authors also thank the EUMETSAT for the support with the Visiting Scientist Activities funded by the H-SAF Consortium.

Acknowledgments: The authors acknowledge the CHUVA Project (FAPESP Grant 2009/15235-8) and H-SAF Consortium for the data provided for this study. Additionally, the authors acknowledge the infrastructure provided by the Institute of Atmospheric Sciences and Climate (CNR-ISAC), the Italian Civil Protection Department (DPC) and National Institute for Space Research (INPE).

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

References

1. 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]
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. Meteorol. Soc.* **2014**, *95*, 701–722. [CrossRef]
3. Mugnai, A.; Smith, E.A.; Tripoli, G.J.; Mugnai, A.; Smith, E.A.; Tripoli, G.J. Foundations for Statistical–Physical Precipitation Retrieval from Passive Microwave Satellite Measurements. Part II: Emission-Source and Generalized Weighting-Function Properties of a Time-dependent Cloud-Radiation Model. *J. Appl. Meteorol.* **1993**, *32*, 17–39. [CrossRef]
4. Smith, E.; Bauer, P.; Marzano, F.; Kummerow, C.; McKague, D.; Mugnai, A.; Panegrossi, G. Intercomparison of microwave radiative transfer models for precipitating clouds. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 541–549. [CrossRef]
5. 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]
6. 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]

7. 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]
8. Bringi, V.; Chandrasekar, V. *Polarimetric Doppler Radar: Principles and Applications*; Cambridge University Press: Cambridge, UK, 2001; p. 636.
9. Machado, L.A.T.; Silva Dias, M.A.F.; Morales, C.; Fisch, G.; Vila, D.; Albrecht, R.; Goodman, S.J.; Calheiros, A.J.P.; Biscaro, T.; Kummerow, C.; et al. The CHUVA project: How does convection vary across Brazil? *Bull. Am. Meteorol. Soc.* **2014**, *95*, 1365–1380. [CrossRef]
10. Calheiros, A.J.P.; Machado, L.A.T. Cloud and rain liquid water statistics in the CHUVA campaign. *Atmos. Res.* **2014**, *144*, 126–140. [CrossRef]
11. Nunes, A.M.P.; Silva Dias, M.A.F.; Anselmo, E.M.; Morales, C.A. Severe Convection Features in the Amazon Basin: A TRMM-Based 15-Year Evaluation. *Front. Earth Sci.* **2016**, *4*, 1–14. [CrossRef]
12. Carvalho, L.M.V.; Jones, C.; Silva, A.E.; Liebmann, B.; Silva Dias, P.L. The South American Monsoon System and the 1970s climate transition. *Int. J. Climatol.* **2011**, *31*, 1248–1256. [CrossRef]
13. 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]
14. Bech, J.; Codina, B.; Lorente, J.; Bebbington, D. The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient. *J. Atmos. Ocean. Technol.* **2003**, *20*, 845–855. [CrossRef]
15. 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]
16. Vulpiani, G.; Tabary, P.; Du Chatelet, J.P.; 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]
17. Gematronik. *Dual-Polarization Weather Radar Handbook*, 2nd ed.; Bringi, V.N., Thurai, M., Eds.; Selex-SI Gematronik: Neuss, Germany, 2007.
18. Marshall, J.S.; Palmer, W.M. The size distribution of raindrops. *Q. J. R. Meteorol. Soc.* **1950**, *76*, 16–36. [CrossRef]
19. Schneebeli, M.; Sakuragi, J.; Biscaro, T.; Angelis, C.F.; Carvalho Da Costa, I.; Morales, C.; Baldini, L.; MacHado, L.A. Polarimetric X-band weather radar measurements in the tropics: Radome and rain attenuation correction. *Atmos. Meas. Tech.* **2012**, *5*, 2183–2199. [CrossRef]
20. 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]
21. Mugnai, A.; Smith, E.A.; Tripoli, G.J.; Bizzarri, B.; Casella, D.; Dietrich, S.; Di Paola, F.; Panegrossi, G.; Sanò, P. CDRD and PNPR satellite passive microwave precipitation retrieval algorithms: EuroTRMM/EURAINSAT origins and H-SAF operations. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 887–912. [CrossRef]
22. 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]
23. Casella, D.; Panegrossi, G.; Sano, P.; Dietrich, S.; Mugnai, A.; Smith, E.A.; Tripoli, G.J.; Formenton, M.; Paola, F.D.; Leung, W.Y. Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive microwave measurements: Part 2. Overcoming database profile selection ambiguity by consideration of meteorological control on microphysics. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4650–4671. [CrossRef]
24. Sano, P.; Casella, D.; Mugnai, A.; Schiavon, G.; Smith, E.A.; Tripoli, G.J. Transitioning from CRD to CDRD in Bayesian retrieval of rainfall and Physics from satellite passive microwave measurements: Part 3—Atmospheric Identification of optimal meteorological tags. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 1185–1208. [CrossRef]
25. Sanò, P.; Panegrossi, G.; Casella, D.; Di Paola, F.; Milani, L.; Mugnai, A.; Petracca, M.; Dietrich, S. The Passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for AMSU/MHS observations: Description and application to European case studies. *Atmos. Meas. Tech.* **2015**, *8*, 837–857. [CrossRef]

26. Sanò, P.; Panegrossi, G.; Casella, D.; Marra, A.C.; Di Paola, F.; Dietrich, S. The new Passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for the cross-track scanning ATMS radiometer: Description and verification study over Europe and Africa using GPM and TRMM spaceborne radars. *Atmos. Meas. Tech.* **2016**, *9*, 5441–5460. [CrossRef]
27. 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]
28. Chen, F.; Staelin, D. AIRS/AMSU/HSB precipitation estimates. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 410–417. [CrossRef]
29. Elsaesser, G.S.; Kummerow, C.D. The sensitivity of rainfall estimation to error assumptions in a Bayesian passive microwave retrieval algorithm. *J. Appl. Meteorol. Climatol.* **2015**, *54*, 408–422. [CrossRef]

remote sensing

MDPI

Article

Performance of TRMM TMPA 3B42 V7 in Replicating Daily Rainfall and Regional Rainfall Regimes in the Amazon Basin (1998–2013)

Véronique Michot [1,*], Daniel Vila [2], Damien Arvor [1], Thomas Corpetti [1], Josyane Ronchail [3,4], Beatriz M. Funatsu [5] and Vincent Dubreuil [1]

1 French National Center for Scientific Research (CNRS), LETG, Université Rennes, UMR 6554 F-35000, France; damien.arvor@univ-rennes2.fr (D.A.); thomas.corpetti@univ-rennes2.fr (T.C.); vincent.dubreuil@univ-rennes2.fr (V.D.)

2 Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), Instituto Nacional de Pesquisas Espaciais (INPE), São Jos é dos Campos, SP 12227-010, Brazil; daniel.vila@cptec.inpe.br

3 University Paris Diderot, Sorbonne Paris Cité Case Courrier 7001, 75205 Paris CEDEX 13, France; josyane.ronchail@locean-ipsl.upmc.fr

4 LOCEAN, Institut Pierre Simon Laplace 4 place Jussieu, 75252 Paris CEDEX 05, France

5 CNRS, Université de Nantes, UMR LETG, Campus du Tertre BP 81227, 44312 Nantes CEDEX 3, France; beatriz.funatsu@univ-nantes.fr

* Correspondence: veronique.michot.pro@gmail.com; Tel.: +33-299-141-847

Received: 4 October 2018; Accepted: 21 November 2018; Published: 24 November 2018

Abstract: Knowledge and studies on precipitation in the Amazon Basin (AB) are determinant for environmental aspects such as hydrology, ecology, as well as for social aspects like agriculture, food security, or health issues. Availability of rainfall data at high spatio-temporal resolution is thus crucial for these purposes. Remote sensing techniques provide extensive spatial coverage compared to ground-based rainfall data but it is imperative to assess the quality of the estimates. Previous studies underline at regional scale in the AB, and for some years, the efficiency of the Tropical Rainfall Measurement Mission (TRMM) 3B42 Version 7 (V7) (hereafter 3B42) daily product data, to provide a good view of the rainfall time variability which is important to understand the impacts of El Nino Southern Oscilation. Then our study aims to enhance the knowledge about the quality of this product on the entire AB and provide a useful understanding about his capacity to reproduce the annual rainfall regimes. For that purpose we compared 3B42 against 205 quality-controlled rain gauge measurements for the period from March 1998 to July 2013, with the aim to know whether 3B42 is reliable for climate studies. Analysis of quantitative (Bias, Relative RMSE) and categorical statistics (POD, FAR) for the whole period show a more accurate spatial distribution of mean daily rainfall estimations in the lowlands than in the Andean regions. In the latter, the location of a rain gauge and its exposure seem to be more relevant to explain mismatches with 3B42 rather than its elevation. In general, a good agreement is observed between rain gauge derived regimes and those from 3B42; however, performance is better in the rainy period. Finally, an original way to validate the estimations is by taking into account the interannual variability of rainfall regimes (i.e., the presence of sub-regimes): four sub-regimes in the northeast AB defined from rain gauges and 3B42 were found to be in good agreement. Furthermore, this work examined whether TRMM 3B42 V7 rainfall estimates for all the grid points in the AB, outgoing longwave radiation (OLR) and water vapor flux patterns are consistent in the northeast of AB.

Keywords: daily rainfall estimations; TRMM 3B42 v7; rain gauges; Amazon Basin; regional rainfall regimes; regional rainfall sub-regimes

1. Introduction

The knowledge of climatic patterns and the monitoring of precipitation are essential for the management of water resources, agriculture, transportation, well-being, health monitoring, erosion, and fire [1–4]. This knowledge, in turn, depends on the availability of rainfall data with adequate spatial and time density. Climatic studies usually rely on long time series (e.g., at least 30 years) to carry out trend analysis. For this reason, and because of their high accuracy, long time series observations from rain gauges are preferred to remote-sensing based estimates. Rain gauges are, however, scarce and unevenly distributed in the Amazon Basin (AB) [2,5–8], and only provide local information that is hard to extrapolate, especially in the Andes that represent approximately 10% (more than 600,000 km^2) of the AB superficies. Indeed, extrapolation could not be representative of the high spatial variability of rainfall in the AB. In contrast, estimated data from remote sensing are attractive as they allow enhancing the spatial information. Nevertheless, satellite products are subject to measurement errors that depend on the geographic context, sensor calibration, and/or algorithm, then they need to be controlled and calibrated, especially with ground data.

The TRMM 3B42 V7 product (afterward 3B42) provide rainfall estimates in the tropics covering nearly three decades (1998–2015). Unlike TRMM 3B42 Real Time product, 3B42 is improved by rain gauges adjustment [9,10]. Despite this adjustment, Delahaye [8] showed many qualitative (hit or miss of rainy events) and quantitative estimation errors in 3B42 version 6 in the Brazilian Legal Amazon. These errors partially arise due to a lower capacity of multi-satellite based algorithms—such as those used by TRMM—to detect warm top stratiform cloud systems [11–13] particularly in mountainous regions, where precipitations tend to be underestimated [14,15]. Still, 3B42 generally shows high performance, if not the best, compared to other remote sensing products [8,13,15,16]. The adequacy of rainfall estimates to represent observations is, however, regionally dependent.

The quality of the TRMM products the most recent version of 3B42 (version 7) has been assessed for several tropical regions, some of them focusing on sub-regions of the AB. In the southern tropical Andes [17,18] TRMM 3B42 and TRMM-PR 2A25 (not rainfall adjusted) provide a useful information to characterized the location of the orographic rainfall intensification in the Andean slopes known as rainfall hotspots (more than 4000 mm per year) in the foothills of the Andes. These works also showed that estimations tend to underestimate rainfall by 35% in the hotspots and until 80% in the highlands. In addition, Espinoza et al. [18] and Paccini and al. [19], using TRMM-PR 2A25 and TRMM 3B42 v7 highlight the capacity of 3B42 to capture the intra-seasonal variability of rainfall, in different Andean locations. In lowlands, in the Rio Negro catchment, [20] and in the Amazon Brazilian river basin [21] TRMM 3B42 version 6 tends to underestimates rainfall. Zubieta et al. [22] and Zulkafli et al. [12], showed overestimation of precipitations in the Amazon basin of Ecuador, above all during the dry season. Furthermore, Zubieta et al. [22] concluded that TRMM data are not totally reliable as they do not entirely reproduce the seasonal differences of the Marañón and the Ucayali subcatchments and tend to underestimates rainfall during the dry season. However we do not have yet, an overall assessment of rainfall estimates, in particular taking into account the time accuracy, such the rainfall regime and its changes from a year to the next, for the whole Amazon Basin (AB) and for several decades. The purpose of this study is thus twofold: (i) assess the spatial and time quality of 3B42 in the whole AB, and (ii) assess its ability to reproduce rainfall regimes at several time and spatial scales, by comparing rain gauge data against 3B42 gridded values.

The paper is organized as follows: in Section 2, we present the study area, the data and the methodology. The results of 3B42 performance are presented in three main parts in Section 3: first, quantitative and qualitative statistical tools are applied to daily values for the whole period 1998–2013. Secondly, we address the capacity of 3B42 to reproduce the rainfall regimes of the different regions of the AB and their time variability which makes it possible to define sub-regimes. Finally, we examine the coherence of 3B42 derived rainfall patterns with outgoing longwave radiation (OLR) and try to explain them using water vapor fluxes over South America. Conclusions and final remarks are presented in Section 4.

2. Data and Methodology

2.1. Study Area

The AB is located in the intertropical zone, between 6° N and 20.5° S and between 48.5° W and 80.5° W (Figures 1 and 2). The AB presents strong topographic contrasts between highlands in the Andean mountains in its western and southwestern parts and lowlands in its central, northern, and eastern parts. The South American Monsoon System (SAMS) [23–26] is the most important climatic feature affecting the region. During the rainy season, typically spanning from October to March, the South American continent warms up and generates the advection of humidity from the northern tropical Atlantic to the AB and until the La Plata basin by the low level jet (LLJ; Figure 1). This increase in humidity in the La Plata basin helps support the South Atlantic Convergence Zone (SACZ) and its associated heavy rainfall. Conversely, from April to September the SAMS moves into its dry phase in most of the AB. The large latitudinal amplitude and the variety of topographical characteristics give rise to a number of rainfall regimes in the AB poorer [27–29], among others]: equatorial in the west and along the Negro Basin (Figure 1); with a dry and a rainy season in the tropical regions of the basin. Over the Atlantic Ocean the convergence of northern and southern trade winds forms the Intertropical Convergence Zone (ITCZ; Figure 1), characterized by strong convection. The equatorial and tropical Atlantic also warm up at the end of the austral summer leading to the southward migration of the ITCZ, and the triggering of the rainy season in the northeast AB.

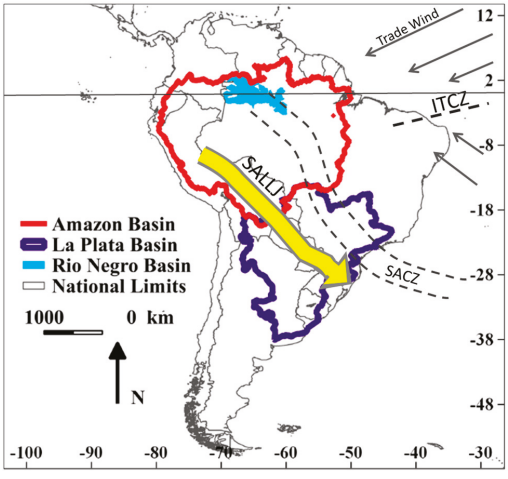


Figure 1. Pattern of the Intertropical Convergence Zone (ITCZ), South Atlantic Convergence Zone (SACZ), South American Low Level Jet (SALLJ) core and northeast and southeast wind trades during the South American Monsoon System from December to February (DJF). Source: V. Michot.

2.2. Observed Precipitations: Rain Gauge Data

The network of observed daily data used as a reference to validate 3B42 estimated data consists of 205 rain gauges (geographically distributed as shown in Figure 2) that have less than 20% missing values and have been quality controlled [30,31]. However, the lack of rain gauges metadata avoid to provide information like the kind device, and then know their accuracy, which could be of great interest when comparing observed and estimated data.

The study covers the period from March 1998 to July 2013. These data were obtained from the National Water Agency (ANA) and the National Meteorological Institute (INMET) in Brazil, the National Meteorological and Hydrological Institute (INAMHI) in Ecuador, the Hydrological

Meteorological and Environmental Studies Institute (IDEAM) in Colombia, and the National Hydrological and Meteorological Service (SENAMHI) in Peru and Bolivia. Unfortunately, no such data could be collected in Venezuela.

2.3. Estimated Precipitations: TRMM TMPA 3B42 Version 7 Daily Product

The 3B42 daily product is computed using the TRMM Multisatellite Precipitation Analysis (TMPA) algorithm, developed by the NASA Goddard Space Flight Center [32]. It consists in a gridded precipitation product with a spatial resolution of 0.25° × 0.25° (approximately 25 × 25 km near the equator) and is available between 50° N and 50° S. This product is gauge-adjusted, but the independence between estimated and observed data is assumed because rain gauge network used in 3B42 comes from monthly estimates from the Global Precipitation Climatology Centre product, which only have few gauges in the AB. The TMPA algorithm combines multiple independent precipitation estimates from multiple instruments on board: the Advanced Microwave Scanning Radiometer for Earth Observing Systems (AMSR-E), Special Sensor Microwave Imager (SSMI), Special Sensor Microwave Imager/Sounder (SSMIS), Advanced Microwave Sounding Unit (AMSU), and Microwave Humidity Sounder (MHS), each calibrated to the TRMM Combined Instrument (TCI). Coverage gaps in space and time are completed with merged and microwave-adjusted geo-infrared (IR) data.

2.4. Outgoing Longwave Radiation and Water Vapor Flux

The satellite measures of OLR were obtained from the National Center for Atmospheric Research of the National Ocean and Atmospheric Administration (NCAR/NOAA) [33] and water vapor flux reanalysis [34] is from the National Center for Environmental Prediction (NCEP{XE "NCEP-NCAR" \t "*National Center for Environnemental Prediction—National Center for Atmospheric Research*"}). The water vapor flux is computed from the specific humidity and the horizontal wind (zonal -u-) from the ground to 300 h Pa [35] as:

$$F_W = \int q\, V\, p/g \tag{1}$$

where q is the humidity, V is the wind vector, p is the pressure, and g is the acceleration due to gravity.

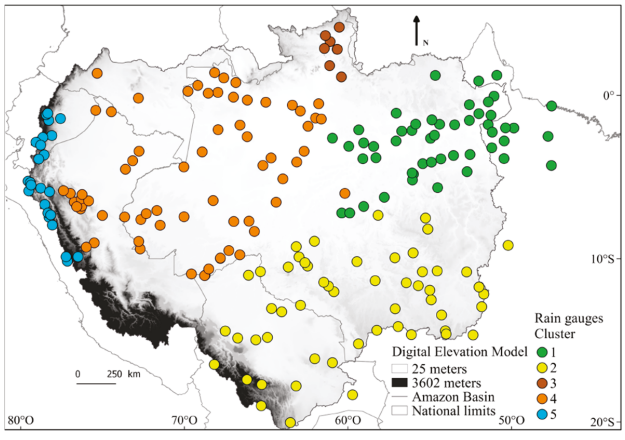


Figure 2. Rain gauge network used as ground reference (205 stations with daily data from 1998 to 2013). The colors of the stations and the labels allow identifying the climatic regions with similar rainfall regimes [33,36].

2.5. Intercomparison Methodology

2.5.1. Daily Rainfall Values

Because of the impossibility to extrapolate the scarce rain gauge network, the comparison between observed and estimated data is based on a point-to-pixel approach, using the pixel that contains the rain gauge (hereafter the 3B42 pixel). This approach was already used in the AB in previous studies and gave accurate results [21,36]. Furthermore, Demirtas et al. showed that quantitative precipitation forecast verification with either a grid to grid or grid to point approach give close results.

The quantitative statistics used in this work are the mean daily rainfall, the Bias Equation (1) and the relative root-mean-square error (relative RMSE, Equation (2)) while the categorical statistics are the probability of detection (POD, Equation (3)) and the False Alarm Ratio (FAR, Equation (4)). These statistics are defined as:

$$Bias = 100\frac{\sum_{i=1}^{N}(Pe,i - Po,\ i)}{\sum_{i=1}^{N} Po,\ i} \tag{2}$$

$$Relative\ RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Po,\ i - Pe,i)^2}{n}}/\overline{Po} \tag{3}$$

$$POD = \frac{H}{H+M} \tag{4}$$

$$FAR = \frac{F}{H+F} \tag{5}$$

where Pe, I is the estimate and Po, I the observation for the i-th day; N is the total number of days, $/\overline{Po}$ is the mean observed precipitation in the whole period, H (hit) is a precipitation event observed by the rain gauge and also detected by 3B42; M (miss) is a precipitation event observed by the rain gauge but not detected by 3B42; F (false alarm) is a precipitation event detected by 3B42 but not observed by the rain gauge. The bias indicates the overestimation or the underestimation by 3B42 in percentage. The relative RMSE gives an average of the error of 3B42, in millimeter (mm). The POD measures the number of rainfall events correctly detected by the estimated product with values ranging between 1 (a perfect score) and 0. The FAR measures the fraction of wrong events detected by 3B42 with values ranging between 1 (the worst score) and 0. The threshold for determining the occurrence of a rain event is set here at 0.1mm in order to avoid the difference in measurement resolution between the two datasets. Indeed, the data from the rain gauges have a resolution of one decimal while 3B42 has several decimals. It is therefore necessary to take it into account in order not to bias the FAR.

In order to assess how well the estimated product reproduces the diversity of rainfall regimes within the AB, the performance of 3B42 is assessed using the standard statistical methods above mentioned at regional scale and for each month of the year. Results of the comparisons are examined, whenever relevant, in the context of the geographical characteristics of the stations (elevation, windward or leeward position).

2.5.2. Rainfall Regimes

In addition to the time-averaged, global 3B42 quality assessment, observed and estimated regional mean annual rainfall and regimes have been computed for five regions, defined using a spectral clustering method [37,38], within the AB (Figure 2). The main principle of spectral clustering is to represent all rain gauges in separate nodes of a connected graph whose vertexes express the similarity between two nodes. The spectral analysis of this graph enables to isolate its main consistent groups. To compute the connection between two nodes into this graph, the basic solution consists in computing a simple Euclidian distance. However, in order to estimate clusters separated in a non-linear way, we exploit the kernel trick. The idea consists in projecting data in another space than the usual one (represented by multi-variate vectors where each component is a value of precipitation) where the separation between clusters is linear. Under some specific properties (see Camps-Valls et Bruzzone [38]

for a complete theory of kernels), this projection can simply be done by changing the way one computes the connection between nodes. In practice, this is done using a Gaussian kernel where the connection between two rain gauges x_1 and x_2 is:

$$K(x_1, x_2) = e^{-\frac{\|x_1 - x_2\|}{\sigma^2}} \quad (6)$$

where σ is a parameter to fix. It has been proven that this kernel enables to efficiently separate highly non-linear clusters.

The determination of the optimal number of clusters is an open problem for which no sound solution exists at the moment. In this study, we rely on the intra/inter inertia. More precisely, a reliable clustering should reveal both homogeneity inside clusters (all stations of the same group are similar) and heterogeneity between averaged clusters (all clusters represent different groups). Therefore, the ratio between the inertia among (averaged) clusters and the internal inertia (sum on inertia inside all groups) should be maximal. The ratio find here is reached when defining 12 clusters. Afterward, small clusters (less than five rain gauges) are gathered with the nearest cluster (taking into account the spatial and interclass variance). The resulting regionalization in Figure 2 is consistent with former studies, showing for example the separation between tropical and equatorial regions and between highlands and lowlands [27,29,39,40]. Finally, these regions, have similar rainfall regimes, and were named according to their relative position within the AB—northeast, south, north, center, and west—as shown in Figure 2.

Then, sub-regimes for each region, based on spectral clustering approach, were also defined in order to assess the interannual variability of regional rainfall patterns. We obtained, for each region, several groups of years—or, clusters—presenting the same type of rainfall sub-regime. Next, based on these clusters rainfall sub-regimes were computed using 3B42. Notice that different clusters not necessarily include the same number of years

Finally, water vapor fluxes and outgoing longwave radiation (OLR) anomalies were compared to precipitation patterns derived from 3B42 all grid point in order (1) to confirm the precipitation patterns and (2) to give some explanations to the observed anomalies. Indeed, OLR is commonly used as proxy for convection in the Tropical region and water vapor fluxes in the AB play an important role in the precipitation variability, and OLR is commonly used as proxy for convection in the Tropical region.

3. Results and Discussion

3.1. Comparison between Points and Pixels at Annual and Monthly Time Scales

At large scale, for the whole AB, 3B42 reproduces well the spatial distribution of mean daily rainfall (Figure 3a,b), with the highest rates of rainfall near the equator while rainfall decreases southward and northward at tropical latitudes. The lowest values are in the Andes, as shown previously. However, 3B42 tends to overestimate rainfall in the lowlands. In the mountains the results vary from one region to another. Indeed, differences can be observed between the northern, central and southern parts of the Andes. In the central part of the Andes, the underestimation is stronger than in the northern part, and rainfall is overestimated in the southern part of the Andes.

The other indicators also depict these differences between lowlands and highlands and within the highlands. The most relevant is the bias (Figure 4a) which shows (a) a null or positive bias in most stations (until +161%), meaning that the estimated values are higher than the observations, and (b) extreme values in the Andean mountains [−45.5; 40] and [+100; +161]% and some exceptions in the northeast and south regions [−40; 35] and [+10; 100]%. In the Andes the bias are positive (until +161%) in most stations of Ecuador (northern Andes), which is coherent with Zubieta et al. [22] and Zulkafli et al. [12] and Bolivia (southern Andes), while the bias is generally negative (until −45.5%) in Peru (Central Andes). This contrast between northern and central Andes was also observed by Zulkafli et al. [12] while the heterogeneity of the bias was found by Salio et al. [13] for the Brazilian

and Bolivian Amazon. Two main hypotheses can be proposed to explain these bias. The first is the location and the leeward or windward exposition of the rain gauges in each region, this point will be discussed later. However, as it is unlikely that all the stations are located leeward or windward, another possibility is that the type of rain gauge differs between the countries and/or during the time series, and have different calibration. Indeed, some rain gauges, those using tipping bucket, can have less precision due to the time to fill the bucket and then to make it tip and register the amount of precipitation during light rainfall. Furthermore, a high evaporation can also cause a lower accuracy of the rain gauge by avoiding a good tipping of the bucket [41,42], because of a same order of magnitude of the precipitation and evaporation rate. Unfortunately, we do not have such information about the rain gauge stations.

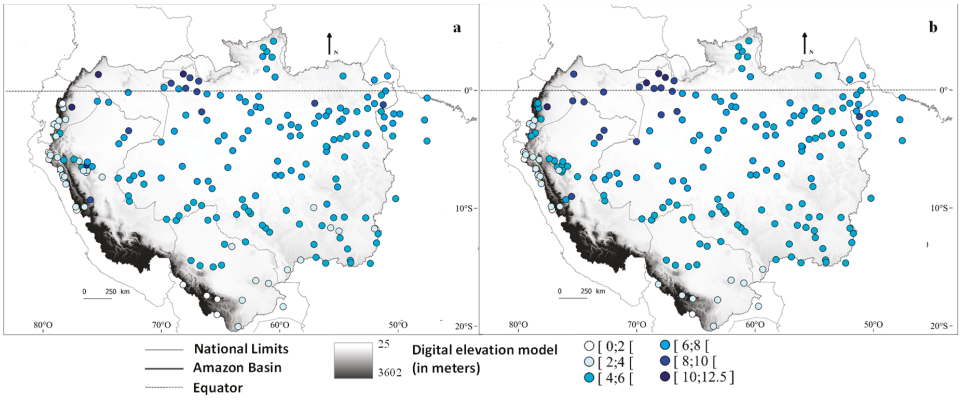


Figure 3. Spatial distribution of (**a**) mean daily rainfall per rain gauge (**b**) mean daily rainfall in the nearest pixel 3B42 of the rain gauge. Both are expressed in millimeters per day. The average is computed on the 1998–2013 period.

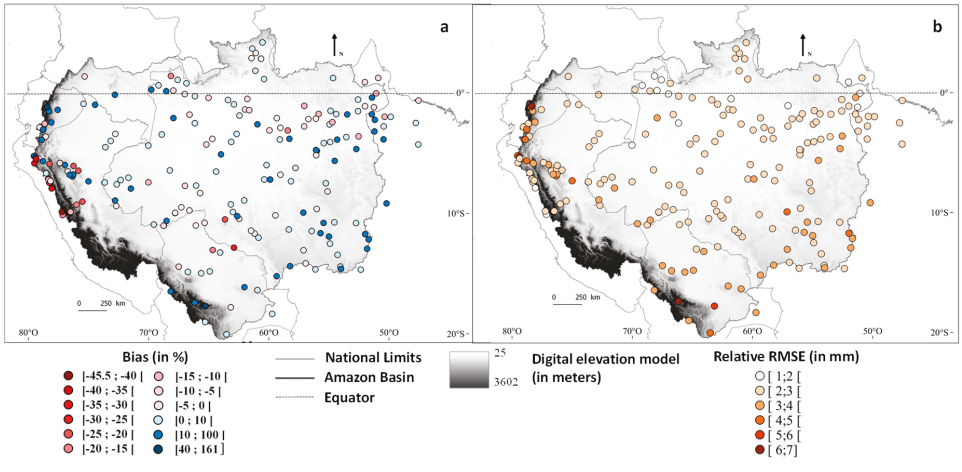


Figure 4. Spatial distribution of (**a**) 3B42 Bias in percent between estimated and observed rainfall, and (**b**) 3B42 Relative RMSE, in millimeters per day. Both are calculated from 2000 to 2013 at daily time scale.

The RMSE (Figure 4b) for most stations was found to be smaller than 3 mm/day, but in some Andean stations and in the southern tropics the RMSE reached values as high as 7 mm/day.

Better results are also observed in the lowlands with the categorical statistics. Indeed, the POD values indicate a good detection (POD > 0.7) of rainy events by the estimated product along a diagonal from northwestern to southeastern Amazon, i.e., along the mean position of the South Atlantic Convergence Zone that produces heavy rainfall in summer (Figure 5a). There is a correct detection of rainy events in the central and southern Andes. The worst values are observed in the northern part of the Andes, in the Bolivian lowlands, and northeast AB.

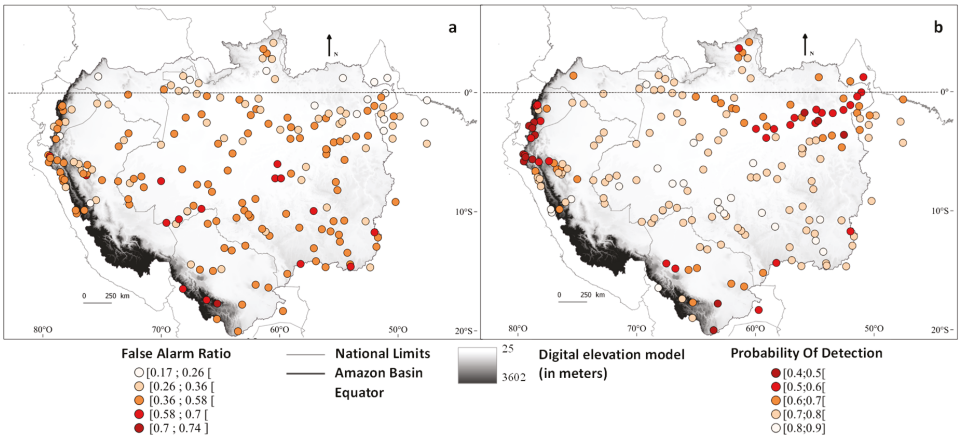


Figure 5. Spatial distribution of (**a**) the probability of detection and (**b**) the false alarm ratio. Both are calculated from 2000 to 2013 at daily time scale.

The false detection of rainy events (FAR) shows on the one hand the Andes and the south of the AB, both of which with the worst results, and on the other, the equatorial and northern locations that show pretty good results (Figure 5b). Globally the results are good with values mainly below 0.58, and there are not clear intra-Andean differences, as observed with other indicators.

Furthermore, as rainfall changes during the year, we computed the Bias, the relative RMSE, the POD and the FAR (Figure 6) at monthly time scale. The results show a contrast between the rainy and the dry seasons, with a higher occurrence of low POD (until 0) and high FAR (until 1), a higher bias (until 300%) and RMSE (until 24 mm) during the driest months. However, since it exists a large range of rainfall regimes in the AB, these time differences could represent only a part of the intra-annual variability, then, a further spatial analysis is needed to better known if that variability is regionally dependent.

3.2. Regional and Time Analysis of TRMM 3B42 V7 Performance

At large scale, for the whole AB, 3B42 reproduces well the spatial distribution of mean daily rainfall (Figure 3a,b), with the highest rates of rainfall near the equator while rainfall decreases toward the tropics. Since the performance of 3B42 varies across the AB the statistical tools are computed at regional and annual time scale. Moreover, as in the previous section they are now computed month by month to examine whether there are any seasonal differences. The regions resulting from the spectral clustering analysis (Figure 2) may be divided into two groups: one where rainy and dry seasons alternate (Figure 7, regions northeast, south, north), and the other without pronounced dry or rainy season (Figure 7, center, west).

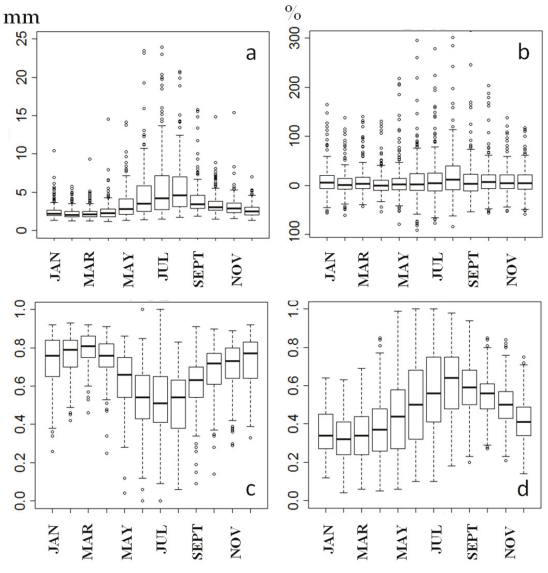


Figure 6. (**a**) Relative RMSE (in mm), (**b**) bias (in mm), (**c**) POD, and (**d**) FAR of 3B42, at daily time scale for each month, calculated from 2000 to 2013 at daily time scale.

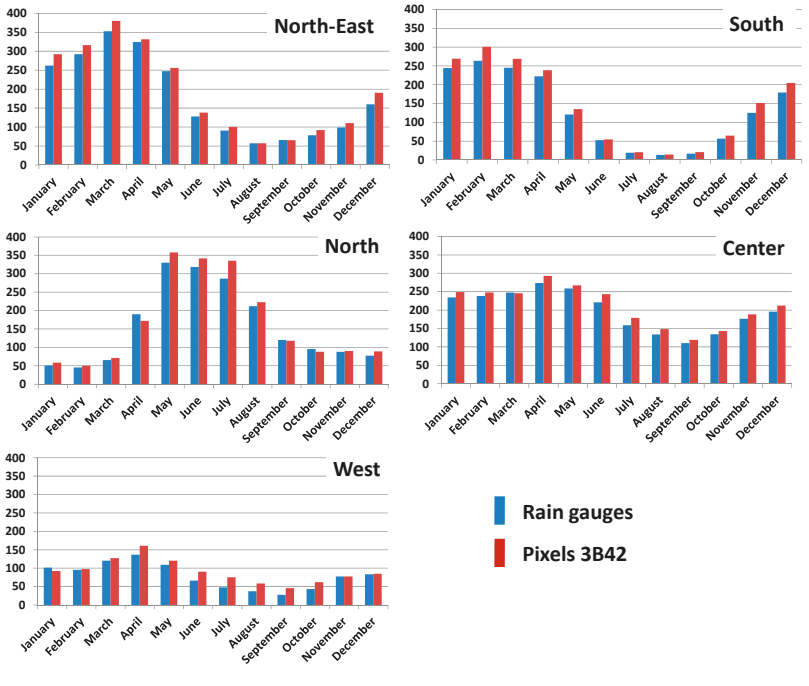


Figure 7. Annual rainfall regimes of the different regions defined in Figure 2 for the AB. The *x*-axis represents each month of the year, *y*-axis represents the rainfall in millimeters per month. Blue bars represent the observed rainfall and red ones the estimated rainfall.

3.2.1. Annual Rainfall in the Regions

In addition to the time-averaged, global 3B42 quality assessment, observed and estimated regional mean annual rainfall and regimes have been computed for five regions within the AB. For all regions the mean annual precipitation was overestimated by 3B42 (Table 1). However, this overestimation is weak in the north and center regions, while south and west regions have the worst results with an average of 6% overestimation.

Table 1. Comparison between rain gauge and 3B42 mean annual rainfall (in mm), in each region of the AB. Differences are computed with respect to rain gauge mean.

Region	3B42	Rain Gauge	Difference 3B42-Rain Gauge
Northeast	2202	2140	+3%
South	1635	1536	+6%
North	1893	1873	+1%
Center	2405	2373	+1%
West	1026	967	+6%

3.2.2. Regional Rainfall Regimes

Figure 7 shows a close reproduction, for all regions, of the mean annual rainfall regimes by 3B42 (despite overestimation in the total mean) with accurate distinction between rainy and dry seasons. The overestimation depends on the month and varies regionally as well: in south, north, and northeast regions rainfall is overestimated during the rainiest months, whereas in center and west regions, the strongest overestimation is observed from March to July.

Furthermore, the statistical parameters show, as for the whole Amazon (Figure 6), a contrast between rainy and dry seasons (Figure 8) with a higher occurrence of low POD and high FAR, a higher bias and RMSE during the driest months. However, that contrast presents a spatial dependence: during their respective driest months the worst bias are near by +5% for the north, below 10% for the center and the northeast and between +20 and 25% for the south and the west region. In this later region, a higher bias during the dry season is coherent with the observation of Zubieta et al. [22] in the Ecuador (until +25%) from 2003 to 2009. However, in each region during their driest months the Bias have a large range of values. On the contrary, during their rainiest months, the bias has lower range of value and the results are better in all the region. These poorer and changing results during the regional driest period, as well as the contrast with the rainiest period (which has better results), are observed for all statistical parameters here (Figure 8).

Once again, the type of rain gauge used could, in part, explain the errors and differences between the regions. Indeed, during the dry season, in the case of tipping bucket rain gauges, if the rainfall is heavy and the bucket take time to fill and tilt, a part of the rainfall can be evaporated [41,42] and not recorded while 3B42 registered this rainfall amount. Unfortunately, because of a lack of information we were not able to take it into account and understand how it may influence the results. However, the type of cloud organization in winter and the difference in spatial resolution of the two datasets [13,16] also appeared to make clear a part of the results. During the dry season, the convection is local, sparse and rainy events are scarcer than during the rainiest months especially in the southern Amazon (south of 6° S) and in the eastern equatorial part of the Amazon [43]. Then a plausible explanation would be that because of its spatial resolution, 3B42 is able to detect a rainy event while the rain gauge cannot, but conversely 3B42 may not detect an event if the convection is too shallow. Additionally, because of the low number of rainy events during the driest months, errors are relatively higher than in other seasons. Conversely, good results in the central region may be due to low seasonal rainfall contrast and deep convection. Future works will further investigate this aspect using cloud cover product as Durieux et al. [44,45] who large range of rainfall examined 10 years of three-hourly infrared data from the International Satellite Cloud Climatology Project.

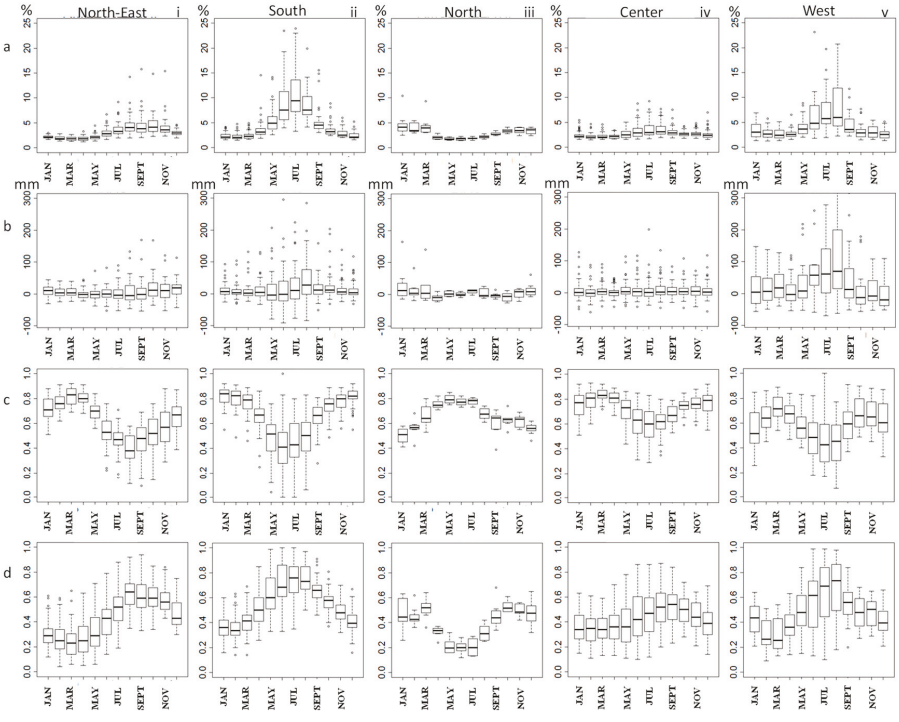


Figure 8. (**a**) Relative RMSE (in mm), (**b**) bias (in mm), (**c**) POD, and (**d**) FAR of 3B42, at daily scale for each month of the year for the five regions: (**i**) the northeast, (**ii**) south, (**iii**) north, (**iv**) center, (**v**) west regions.

The south and west regions present the worst results for all statistical parameters and all months (Figure 8a–d (ii) and (v)), consistent with Figures 4 and 5. In mountainous regions (the western part of the south region and the west region), the accuracy of the estimates tends to decrease with elevation [12,16]. To investigate this issue in more detail the correlation between all the statistical parameters and the elevation was examined. The results show that only POD is slightly but significantly correlated to the elevation (Figure 9c). A possible interpretation would be that at high elevation rainy events are rarer but convective systems are well organized during the summer, which allows a good detection of the events by the satellites and POD scores higher than 0.4. The other parameters do not show a significant relationship between the elevation and the quality of 3B42 estimates, in agreement with Turko who found similar results for 3B42 estimates in the Bolivian Andes.

Some other explanation of the poor results in the Andean region could be the location of the rain gauge itself and its exposure. Indeed, in the western region and in the western part of the southern region, the stations where 3B42 overestimates rainfall are located either at the bottom of a valley or are leeward (Figure 10a,b respectively). One potential interpretation is the alteration of measurements due to evapotranspiration. Indeed, satellites measure rainfall at a higher elevation that rain gauges, then if the evaporation is important, it can avoid the precipitation to hit the ground, like in the case of virga cloud, and then cause a difference of recorded value between observations and estimations [46]. Conversely, due to the orographic effect, some precipitations can form close to the ground by advection of humid air on the hillside well exposed to the wind; in that case, the station is able to catch such precipitations, while satellites are not. Another possibility is that strong winds blow the rainfall and avoid it to enter in the rain gauges when this one are exposed to these winds. However we do not have

this wind exposition information. At topographic scale, the orography also seems to explain some of 3B42 overestimation. Figure 10a presents for instance a case of stations sheltered by a mountain range. Main winds from the northwest are blocked by this rock and form a rainfall hotspot on the northern slope [18].

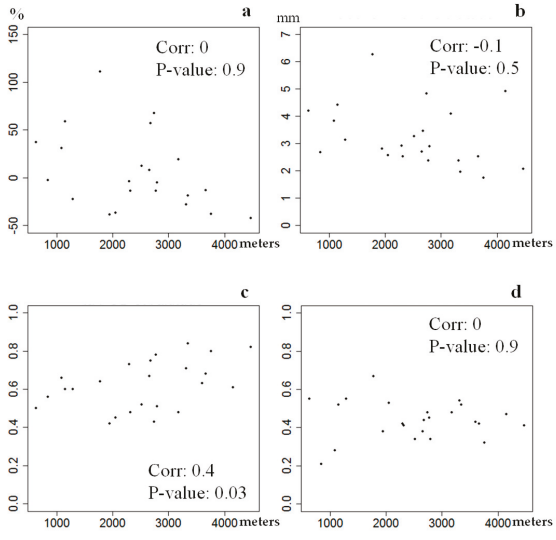


Figure 9. Correlation between elevation and (**a**) bias, in %; (**b**) relative RMSE, in mm; (**c**) POD; and (**d**) FAR in the region west. The *x*-axis represents the elevation (in meters), *y*-axis represents the score of each statistic.

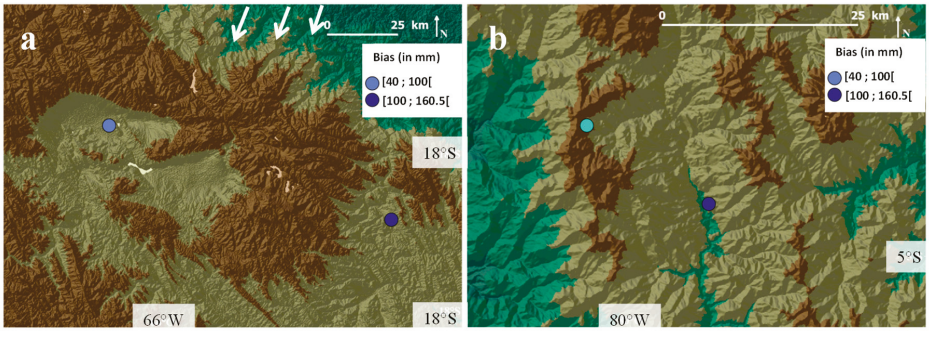


Figure 10. (**a**,**b**) location of rain gauges with shelter situation and overestimation (in mm) by 3B42. The white arrow represent the main wind in the region of Bolivian rainfall hotspot [18].

3.2.3. Regional Rainfall Sub-Regimes in the Northeastern Region of the AB

The interannual regional rainfall variability sub-regimes were first analyzed using the rain gauge network. For each region, between two and four groups of years presenting the same type of rainfall sub-regime were defined using the spectral clustering method. These sub-regimes were assumed to depict the interannual variability of regional precipitation. Hereafter we present the results for the northeastern region (Figure 2), as it contains a relatively large number of rain gauges, and good 3B42 performance.

Table 2 shows the clusters and the years belonging to each cluster and Table 3 rain gauge and 3B42 mean annual rainfall (in mm) for each rainfall sub-regime of the northeastern region. For each rainfall sub-regime, 3B42 tends to overestimate by 2 to 4% the annual precipitation (Table 3), which is coherent with the analysis of the mean regional rainfall regime (Figure 5).

Table 2. Years included in each rainfall sub-regime (Cl1 to Cl4) detected on the rain time series, by spectral clustering method, in the northeastern region of the AB.

Year	Cluster
2002–2003	1
2004–2005	1
2006–2007	1
2011–2012	1
2000–2001	2
2001–2002	2
2009–2010	2
2010–2011	2
2008–2009	3
1998–1999	4
1999–2000	4
2005–2006	4
2007–2008	4

Table 3. Rain gauge and 3B42 mean annual rainfall (in mm) for each rainfall sub-regime of the northeastern region.

Rainfall Sub-Regime	Rain Gauge	3B42	Errors (%) of 3B42
Cl1	2181	2234	2
Cl2	2170	2261	4
Cl3	2545	2607	2
Cl4	2369	2419	2

Despite the overestimation in the mean annual totals, Figure 11 shows that the annual cycles of rainfall of observed and estimated precipitation are similar for the four rainfall sub-regimes in the northeastern region. The seasonal and intra-seasonal inflections of the curve are well reproduced by 3B42. In comparison with a traditional analysis of annual precipitations, this sub-regime approach shows the irregular distribution errors along the year. Furthermore, and in contrast with Figure 7a, the mismatches depend on the rainfall sub-regime and not on a specific month or period of the year. The highest overestimation of annual precipitation in Cl2 (Table 3) appears to be caused by a slight but persistent overestimation all year long (Figure 11), while the strongest overestimation between May and July in Cl3 is smoothed by underestimation during the other months of the year. Despite these differences, the reproduction of precipitation patterns by 3B42 is quite accurate.

3.3. Comparisons between 3B42, Water Vapor Flux, and OLR Spatial Pattern Anomalies

Up to now, a point-to-pixel approach was used in order to validate 3B42 estimates for the AB, however, there still remains the question whether 3B42 delivers spatially coherent patterns if we consider all grid points (i.e., grid points beyond the pixels used for the comparisons with ground station).

To investigate this issue, the monthly anomalies derived from the full grid of 3B42 (hereafter 3B42 grid) were analyzed for the northeastern region. OLR anomalies were then used to confirm the spatial pattern of 3B42 grid anomalies and water vapor flux to try to explain these anomalies. Monthly anomalies of these variables correspond to the difference between the mean value of the years of the rainfall sub-regime and the mean value of the period 1998–2013 and only the significant anomalies

values (higher than 1 standard deviation) are represented in Figures 12 and 13. The precipitation anomalies are normalized, that is, divided by the standard deviation of the whole study period.

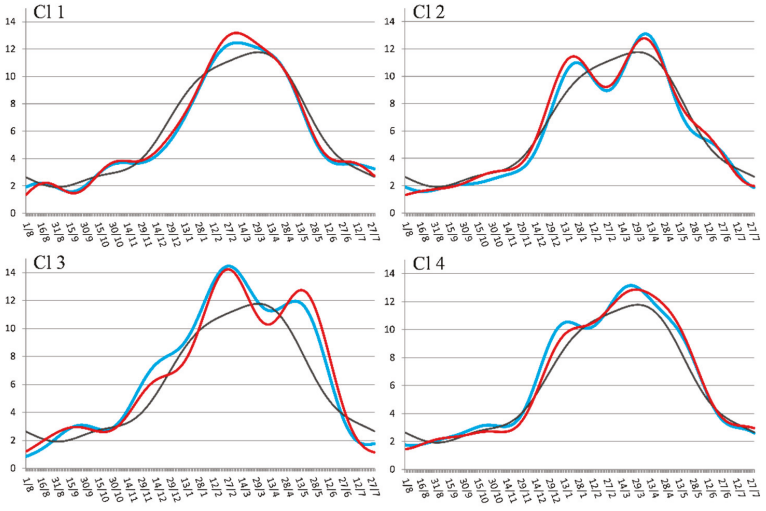


Figure 11. Rainfall sub-regimes Cl1 to Cl4, in the northeastern region of AB from rain gauge (blue line), 3B42 pixel (red line), and mean annual regime of the region from rain gauge (grey line).

The analysis was performed in all the cluster sub-regimes [37], but we focus on the sub-regime Cl1 and Cl4 which are contrasted. Indeed, Cl1 has a shorter peak period than the average cycle of the region (grey line), and Cl4 on the contrary, has a longer rainfall peak (Figure 11). Furthermore, the explanations of the anomalies by the atmospheric circulation are very clear for these two clusters, which are chosen as examples. Cl3 only contains one year which can confuse the results

Figures 12 and 13 show for sub-regimes Cl1 and Cl4, composite maps of normalized monthly anomalies of 3B42 grid and OLR and water vapor flux monthly anomalies. The blue polygon indicates the northeastern region within the AB. A positive (negative) OLR anomaly in warm (cold) color indicates less (more) convection and consequently, a lower (higher) probability of positive (negative) rainfall anomalies. The water vapor flux anomalies represented by vectors indicate if more or less humidity is advected to the AB; a lower (higher) transport of humidity is consistent with lower (higher) convection.

Composites for Cl1 show that 3B42 grid rainfall anomalies are in good agreement with the previous rainfall sub-regimes found at rain gauge and 3B42 pixel. The shorter rainfall peak observed in Figure 11 is in agreement with negative precipitation anomalies in the 3B42 grid in the northeastern region. These anomalies may be homogeneously distributed in the region, as in January and June, or concentrated in certain areas (e.g., in December and May). This highlights the high spatial rainfall variability and the difficulty to take it into account with the rain gauge network only. The anomalies of the 3B42 and OLR composites also are in accordance (Figure 12) and show strong negative anomalies of convection in the northeastern region in December–January and May–June (Figure 12), which is in line with a reduced peak of rainfall in this region occurring usually from December to May. The deficit of precipitation and positive OLR anomalies are consistent with the reduced and divergent water vapor flux from North Tropical Atlantic (NATL). See for instance the lack convergent water vapor fluxes in December–January and May–June over the region, while convergent fluxes are observed from February to April when enhanced precipitation is observed. The excess of rainfall in February–March, during the rainfall peak (Figure 11), appears to be linked with an area with positive rainfall (Figure 11)

close to the coast that is strongly influenced by the water vapor flux from tropical ocean. These fluxes are slightly reinforced in March and may explain the enhanced rainfall during this period of the year. Another feature that appears using the 3B42 grid is the opposed inland and coastal anomalies inside the region (e.g., in March and April, Figure 12).

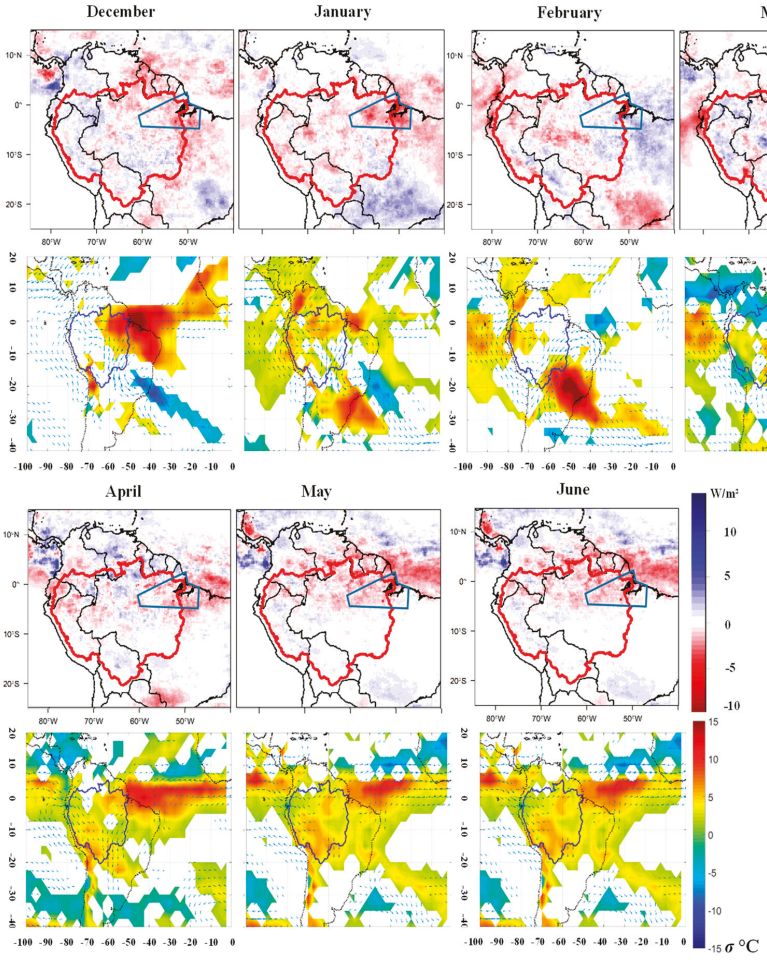


Figure 12. Monthly composite maps of 3B42 grid normalized monthly anomalies (**top**) and OLR and water vapor flux monthly significant anomalies (**bottom**). Anomalies are computed based on cluster Cl1 years in the northeastern region. The blue polygon represents the northeastern region of the AB.

The precipitation anomalies in the 3B42 grid composites of the rainfall sub-regime Cl4 also highlight the importance of the local scale compared to the regional scale. Indeed, during the longer and higher than usual rainy peak of Cl4, with strong surpluses of precipitation in January and from March to June (Figure 11), rainfall anomalies appear in the 3B42 grid composites but with an inhomogeneous spatial distribution (Figure 13). In contrast, in February the slight rainfall deficits in rainfall sub-regime 3B42 pixel (Figure 11) also appears in the 3B42 grid, but the latter have a more spatial homogeneous pattern. The positive rainfall anomaly is connected to negative OLR anomalies (Figure 12) on the northeast of the AB and on the Atlantic ITCZ in January and from March to June.

The convection increase is related to the convergence in the northeastern region between water vapor flux anomalies from the west, the north and the southeast, as well as a weaker than usual export of humidity by the low-level jet to the La Plata Basin. On the contrary, in February even if strong convection is present on the Northeast AB region the water vapor fluxes are reduced which is coherent with the rainfall deficits of 3B42 grid in this region in years where Cl4 dominates.

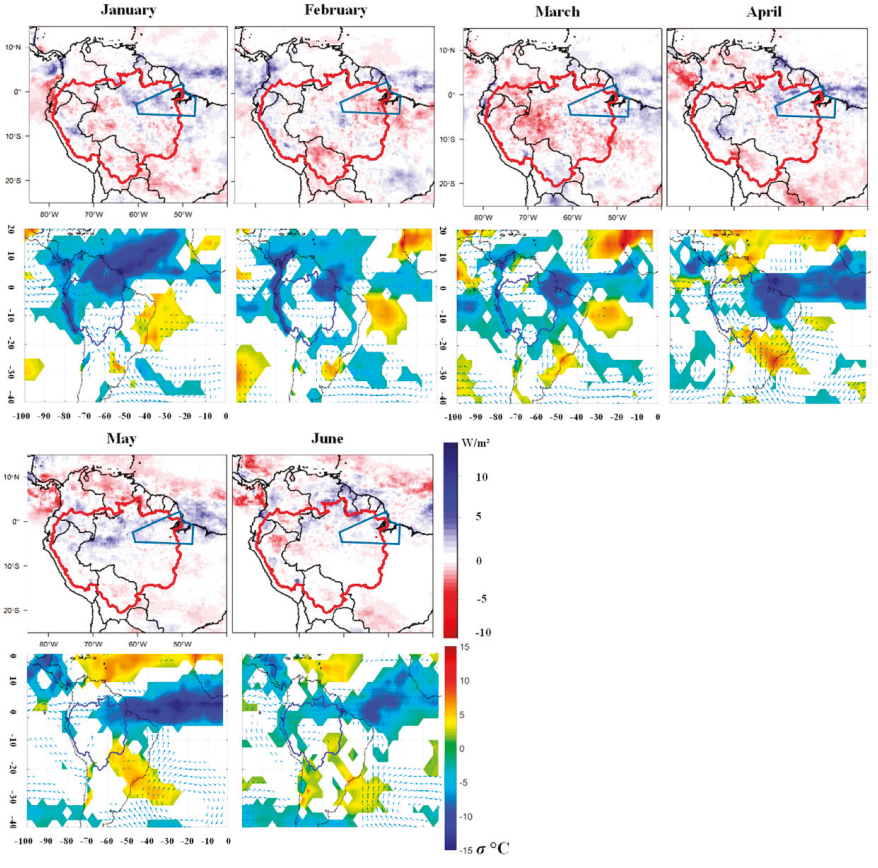


Figure 13. Monthly composite maps of 3B42 grid normalized monthly anomalies (**top**) and OLR and water vapor flux monthly significant anomalies (**bottom**). Anomalies are computed based on cluster Cl4 years in the northeastern region. The blue polygon represents the northeastern region of the AB.

4. Conclusions

The 3B42 daily precipitation product is an incredibly valuable product as it provides spatial rainfall information for the whole AB, where the access to ground observed data is particularly complicated. However, as they are estimated, some bias and errors exist that need to be assessed. The evaluation of the 3B42 quality was done by means of quantitative (mean daily rainfall, bias, relative RMSE) and categorical statistics (POD and FAR), considering daily data in the whole basin. These statistical parameters showed monthly differences with a contrast between the dry and rainy period. This last one presents the better results. In relation to these intra-annual contrasts, we focused our evaluation on regional rainfall regimes and sub-regimes which have not been shown in any previous studies The average bias is +7% (overestimation), but large errors ranging from −45% to

+161% are observed in the Andean regions. 3B42 performs better in the detection of rainy events than in the quantitative estimation of rainfall. Also, as observed in previous studies and as shown by all the statistics, 3B42 performs better in the lowlands of the AB than in mountainous regions of the Andes. However this work highlights differences within the high areas: in the northern (Ecuador) and southern Andes (Bolivia) 3B42 tends to overestimate precipitation, whereas in central Andes (Peru) rainfall is underestimated. Furthermore, no correlation is found between the elevation of the stations and the results of the statistical parameters. Warm clouds are frequently assumed to be a factor of the difference measured between observed and estimated precipitations. However, others phenomena in AB can also cause differences between observed and estimated data. Indeed, overestimation by 3B42 can results from the capacity of satellite to measure precipitation at high elevation, while the evaporation avoids this same precipitation to hit the ground and then, the rain gauges. Underestimations by 3B42 could depend on orography with a low capacity to detect orographic rainfall forming close to the ground on hillside exposed to the main wind, while the rain gauge succeeds in catching it. Furthermore, we assume that further evaluation can be done by considering several threshold values to determine the occurrence of a precipitation event (here, 0.1mm) to check if underestimation depends on such threshold.

The ability of 3B42 to accurately replicate annual rainfall regimes in the different AB regions was also used as a tool to assess its quality. There is generally a good agreement between regional regimes derived either from rain gauge or 3B42. However, in each region, the performance of 3B42 is lower during the drier period than during the rainy period. That could be partially explained by the difference in spatial resolution of the two datasets (the sparse occurrence of rainfall not being measured by the rain gauges but being captured by 3B42). Also, during the drier period, rainy events are scarcer and the relative error may be larger.

Rainfall sub-regimes from the rain gauge and those from the 3B42 pixels were compared and are in good agreement in the northeastern region, despite an overestimation of 2 to 4% of annual rainfall. The seasonal and intra-seasonal behavior is reproduced well. It is worth noting that over or underestimation of 3B42 pixel do not depend on the months or the season, and differ between the rainfall sub-regimes.

Finally, the adequacy between observed and estimated rainfall sub-regimes were compared looking at the whole northeast grid. Generally, both are in good agreement. However, it can be noted that sometimes precipitations anomalies are not homogeneous in the northeastern region, for instance with differences between inland and coastal sectors. That underlines a high spatial rainfall variability that is difficult to take into account with a rain gauge network only, and highlights the importance of spatially refined data such as 3B42. Spatial and time accuracy of 3B42 grid anomalies are coherent with the water vapor flux and OLR anomalies, despite their different resolutions.

At a daily scale, in most parts of AB, 3B42 provides an exhaustive and quite good precipitation dataset. This higher quantity of information with satisfactory quality emphasizes intraregional precipitation distribution, especially between coastal (often influenced by sea breeze front) and inland sectors. Nevertheless, when using 3B42 for the AB, data should be used with caution in the dry season or in mountainous regions. Beyond these aspects, the upcoming data homogenization between 3B42 and the recent Global Precipitation Measurement Mission will provide a longer and improved rainfall estimated time series.

Author Contributions: V.M., D.V., T.C., and D.A. designed the experiments; V.M., T.C., and D.A. performed the calculations; all the authors analyzed the results and wrote the paper.

Funding: This research was funded by a followship of outgoing mobility from Rennes Métropole and the H2020-MSCA-RISE-2015 ODYSSEA project (project reference, 691053).

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

References

1. Meinke, H.; Stone, R.C. Seasonal and Inter-Annual Climate Forecasting: The New Tool for Increasing Preparedness to Climate Variability and Change in Agricultural Planning and Operations. *Clim. Chang.* **2005**, *70*, 221–253. [CrossRef]
2. Arvor, D. Etude par Télédétection de la Dynamique du soja et de L'impact des Précipitations sur les Productions au Mato Grosso (Brésil). Thèse de Doctorat, Université Rennes 2, Rennes, France, 2009.
3. Brondizio, E.S.; Moran, E.F. Human dimensions of climate change: The vulnerability of small farmers in the Amazon. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2008**, *363*, 1803–1809. [CrossRef] [PubMed]
4. Coomes, O.T.; Lapointe, M.; Templeton, M.; List, G. Amazon river flow regime and flood recessional agriculture: Flood stage reversals and risk of annual crop loss. *J. Hydrol.* **2016**, *539*, 214–222. [CrossRef]
5. Liebmann, B.; Allured, D. Daily precipitation grids for South America. *Bull. Am. Meteorol. Soc.* **2005**, *86*, 1567. [CrossRef]
6. Dubreuil, V.; Arvor, D.; Ronchail, J. Potentialités des données TRMM pour la spatialisation des précipitations au Mato Grosso, Brésil. In Proceedings of the XXe Colloque de l'Association Internationale de Climatologie, Carthage, Tunisie, 3–9 September 2007; pp. 210–215.
7. Ronchail, J.; Cochonneau, G.; Molinier, M.; Guyot, J.-L.; De Miranda Chaves, A.G.; Guimarães, V.; de Oliveira, E. Interannual rainfall variability in the Amazon basin and sea-surface temperatures in the equatorial Pacific and the tropical Atlantic Oceans. *Int. J. Climatol.* **2002**, *22*, 1663–1686. [CrossRef]
8. Delahaye, F. Analyse Comparative des Différents Produits Satellitaires D'estimation des Précipitations en Amazonie Brésilienne. Thèse de Doctorat, Université Rennes 2, Rennes, France, 2013.
9. Cai, Y.; Jin, C.; Wang, A.; Guan, D.; Wu, J.; Yuan, F.; Xu, L. Spatio-temporal analysis of the accuracy of tropical multisatellite precipitation analysis 3B42 precipitation data in mid-high latitudes of China. *PLoS ONE* **2015**, *10*, e0120026. [CrossRef] [PubMed]
10. Guo, R.; Liu, Y. Evaluation of Satellite Precipitation Products with Rain Gauge Data at Different Scales: Implications for Hydrological Applications. *Water* **2016**, *8*, 281. [CrossRef]
11. 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]
12. Zulkafli, Z.; Buytaert, W.; Onof, C.; Manz, B.; Tarnavsky, E.; Lavado, W.; Guyot, J.L. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean–Amazon River basins. *J. Hydrometeorol.* **2014**, *15*, 581–592. [CrossRef]
13. Salio, P.; Hobouchian, M.P.; Skabar, Y.G.; Vila, D. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmos. Res.* **2015**, *163*, 146–161. [CrossRef]
14. Dinku, T.; Connor, S.J.; Ceccato, P. Comparison of CMORPH and TRMM-3B42 over Mountainous Regions of Africa and South America. In *Satellite Rainfall Applications for Surface Hydrology*; Gebremichael, M., Hossain, F., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp. 193–204. ISBN 978-90-481-2915-7.
15. Mohd Zad, S.; Zulkafli, Z.; Muharram, F. Satellite Rainfall (TRMM 3B42-V7) Performance Assessment and Adjustment over Pahang River Basin, Malaysia. *Remote Sens.* **2018**, *10*, 388. [CrossRef]
16. 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]
17. Bookhagen, B.; Strecker, M.R. Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes. *Geophys. Res. Lett.* **2008**, *35*, L06403. [CrossRef]
18. Espinoza, J.C.; Chavez, S.; Ronchail, J.; Junquas, C.; Takahashi, K.; Lavado, W. Rainfall hotspots over the southern tropical Andes: Spatial distribution, rainfall intensity, and relations with large-scale atmospheric circulation. *Water Resour. Res.* **2015**, *51*, 3459–3475. [CrossRef]
19. Paccini, L.; Espinoza, J.C.; Ronchail, J.; Segura, H. Intra-seasonal rainfall variability in the Amazon basin related to large-scale circulation patterns: A focus on western Amazon-Andes transition region: Intra-seasonal rainfall variability in western Amazon. *Int. J. Climatol.* **2018**, *38*, 2386–2399. [CrossRef]
20. Getirana, A.C.; Espinoza, J.C.V.; Ronchail, J.; Rotunno Filho, O.C. Assessment of different precipitation datasets and their impacts on the water balance of the Negro River basin. *J. Hydrol.* **2011**, *404*, 304–322. [CrossRef]

21. Buarque, D.C.; de Paiva, R.C.D.; Clarke, R.T.; Mendes, C.A.B. A comparison of Amazon rainfall characteristics derived from TRMM, CMORPH and the Brazilian national rain gauge network. *J. Geophys. Res.* **2011**, *116*. [CrossRef]
22. Zubieta, R.; Getirana, A.; Espinoza, J.C.; Lavado, W. Impacts of satellite-based precipitation datasets on rainfall–runoff modeling of the Western Amazon basin of Peru and Ecuador. *J. Hydrol.* **2015**, *528*, 599–612. [CrossRef]
23. Zhou, J.; Lau, K.M. Does a Monsoon Climate Exist over South America? *J. Clim.* **1998**, *11*, 1020–1040. [CrossRef]
24. Marengo, J.A.; Liebmann, B.; Kousky, V.E.; Filizola, N.P.; Wainer, I.C. Onset and End of the Rainy Season in the Brazilian Amazon Basin. *J. Clim.* **2001**, *14*, 833–852. [CrossRef]
25. Wang, H.; Fu, R. Cross-Equatorial Flow and Seasonal Cycle of Precipitation over South America. *J. Clim.* **2002**, *15*, 1591–1608. [CrossRef]
26. Vera, C.; Baez, J.; Douglas, M.; Emmanuel, C.B.; Marengo, J.; Meitin, J.; Nicolini, M.; Nogues-Paegle, J.; Paegle, J.; Penalba, O.; et al. The South American Low-Level Jet Experiment. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 63–77. [CrossRef]
27. Figueroa, S.N.; Nobre, C.A. Precipitation distribution over central and western tropical South America. *Climanalise* **1990**, *5*, 36–45.
28. Marengo, J.A. Interannual variability of surface climate in the Amazon basin. *Int. J. Climatol.* **1992**, *12*, 853–863. [CrossRef]
29. Espinoza Villar, J.C.; Ronchail, J.; Guyot, J.L.; Cochonneau, G.; Naziano, F.; Lavado, W.; De Oliveira, E.; Pombosa, R.; Vauchel, P. Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *Int. J. Climatol.* **2009**, *29*, 1574–1594. [CrossRef]
30. Michot, V.; Dubreuil, V.; Roncahil, J.; Lucio, P.S. Constitution et analyse critique d'une base de données spatiale pour l'étude des saisons des pluies dans le bassin amazonien. In Proceedings of the ENVIBRAS 2014 Environnement et Géomatique: Approches Comparées France—Brésil, Rennes, France, 12–15 November 2014; pp. 237–244.
31. Michot, V.; Arvor, D.; Ronchail, J.; Corpetti, T.; Jégou, N.; Lucio, P.S.; Dubreuil, V. Validation and reconstruction of rain gauge-based daily time series for the entire Amazon basin. *Theor. Appl. Climatol.* **2018**. under review.
32. Huffman, G.J.; Bolvin, D.T. TRMM and Other Data Precipitation Data Set Documentation 2015. Available online: https://pmm.nasa.gov/sites/default/files/document_files/3B42_3B43_doc_V7.pdf (accessed on 10 August 2015).
33. Liebmann, B.; Smith, A.C. Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 1275–1277.
34. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–471. [CrossRef]
35. Peixoto, J.P.; Oort, A.H. *Physics of Climate*; American Institute of Physics: New York, NY, USA, 1992; 520p, ISBN1 978-0-88318-711-1. ISBN2 978-0-88318-712-8.
36. Scheel, M.L.M.; Rohrer, M.; Huggel, C.; Santos Villar, D.; Silvestre, E.; Huffman, G.J. Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2649–2663. [CrossRef]
37. Michot, V. Analyse Spatiale et Temporelle de la Variabilité des Régimes de Précipitations dans le Bassin Amazonien. Ph.D. Thesis, Université Bretagne-Loire, Rennes 2, Rennes, France, 2017.
38. Camps-Valls, G.; Bruzzone, L. *Kernel Methods for Remote Sensing Data Analysis*; John Wiley & Sons: Chichester, UK, 2009; ISBN 978-0-470-74899-2.
39. Barbosa Santos, E.B. Modelagem estatística e atribuições dos eventos de precipitação extrema na Amazônia brasileira. In *Statistical Modeling and Attributions of Extreme Precipitation Events in the Brazilian Amazon*; Universidade Federal do Rio Grande: Do Norte Natal, Brazil, 2015.
40. Delahaye, F.; Kirstetter, P.-E.; Dubreuil, V.; Machado, L.A.; Vila, D.A.; Clark, R. A consistent gauge database for daily rainfall analysis over the Legal Brazilian Amazon. *J. Hydrol.* **2015**, *527*, 292–304. [CrossRef]
41. Groisman, P.Y.; Legates, D.R. The Accuracy of United States Precipitation Data. *Bull. Am. Meteorol. Soc.* **1994**, *75*, 215–227. [CrossRef]

42. Wang, J.; Fisher, B.L.; Wolff, D.B. Estimating Rain Rates from Tipping-Bucket Rain Gauge Measurements. *J. Atmos. Oceanic Technol.* **2008**, *25*, 43–56. [CrossRef]
43. Fu, R.; Zhu, B.; Dickinson, R.E. How Do Atmosphere and Land Surface Influence Seasonal Changes of Convection in the Tropical Amazon? *J. Clim.* **1999**, *12*, 1306–1321. [CrossRef]
44. Durieux, L. Etude des Relations entre les Caractéristiques Géographiques de la Surface et les Nuages Convectifs dans la Région de L'arc de Déforestation en Amazonie. Ph.D. Thesis, Université de Provence, UFR Des Sciences Géographiques et de L'aménagement, Aix-Marseille, France, 2002.
45. Durieux, L. The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sens. Environ.* **2003**, *86*, 132–140. [CrossRef]
46. Wang, Z.; Sassen, K. Cloud Type and Macrophysical Property Retrieval Using Multiple Remote Sensors. *J. Appl. Meteorol.* **2001**, *40*, 1665–1682. [CrossRef]

remote sensing

MDPI

Technical Note

How Well Can Global Precipitation Measurement (GPM) Capture Hurricanes? Case Study: Hurricane Harvey

Ehsan Omranian [1,*], Hatim O. Sharif [1] and Ahmad A. Tavakoly [2]

[1] Department of Civil and Environmental Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA; hatim.sharif@utsa.edu

[2] Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, 5825 University Research Ct suite 4001, College Park, MD 20740, USA; Ahmad.A.Tavakoly@erdc.dren.mil

* Correspondence: seyedehsan.omranian@utsa.edu; Tel.: +1-210-803-3847

Received: 12 June 2018; Accepted: 18 July 2018; Published: 20 July 2018

Abstract: Hurricanes and other severe coastal storms have become more frequent and destructive during recent years. Hurricane Harvey, one of the most extreme events in recent history, advanced as a category IV storm and brought devastating rainfall to the Houston, TX, region during 25–29 August 2017. It inflicted damage of more than $125 billion to the state of Texas infrastructure and caused multiple fatalities in a very short period of time. Rainfall totals from Harvey during the 5-day period were among the highest ever recorded in the United States. Study of this historical devastating event can lead to better preparation and effective reduction of far-reaching consequences of similar events. Precipitation products based on satellites observations can provide valuable information needed to understand the evolution of such devastating storms. In this study, the ability of recent Integrated Multi-satellitE Retrievals for Global Precipitation Mission (GPM-IMERG) final-run product to capture the magnitudes and spatial ($0.1^\circ \times 0.1^\circ$)/temporal (hourly) patterns of rainfall resulting from hurricane Harvey was evaluated. Hourly gridded rainfall estimates by ground radar (4×4 km) were used as a reference dataset. Basic and probabilistic statistical indices of the satellite rainfall products were examined. The results indicated that the performance of IMERG product was satisfactory in detecting the spatial variability of the storm. It reconstructed precipitation with nearly 62% accuracy, although it systematically under-represented rainfall in coastal areas and over-represented rainfall over the high-intensity regions. Moreover, while the correlation between IMERG and radar products was generally high, it decreased significantly at and around the storm core.

Keywords: hurricane Harvey; GPM satellite; IMERG; tropical storm rainfall; gridded radar precipitation

1. Introduction

Reliable and high-quality precipitation measurements provide indispensable information for water resources investigation and play a substantial role in various applications such as weather prediction and climatic modeling, water resources and hydrologic modeling, global water cycle analysis, and agricultural development planning [1–7]. However, obtaining accurate and authoritative data such as precipitation products has always been a challenging issue for scientists due to the great heterogeneity across uneven spatiotemporal distribution of water resources on both continental and regional scales [8–11]. The amount of precipitation is mainly recorded via three methods: ground observations (gauges), weather radars, and satellite-based products [12,13]. Conventionally, rain gauge networks, as the most straightforward and reliable measurement method, provide the surface precipitation amount with high temporal frequency [8,14,15]. However, uneven and sparse distribution of gauges causes lower spatial coverage of larger watersheds. Moreover, high data latency

in an obstacle to the wide use of gauge networks [11,16,17]. Ground-based weather radars generally estimate precipitation with higher spatiotemporal resolution which can significantly lead to better understanding of storm structure and their real-time monitoring [18–21]. Despite some constraints such as limited spatial coverage, inaccuracy in mountainous regions, range-dependent and random systematic errors; radar precipitation products are currently one of the most trusted sources for precipitation measurements [22–24]. Multi-Satellite precipitation measurements can be potentially the most comprehensive products due to their vast spatial coverage (almost globally), regular and constant temporal measurements, and moderately synchronized performance [25–29]. For nearly two decades, the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) has been providing high quality precipitation records which have been used in various applications [30]. This product is recently integrated with some other semi-global satellite products such as Climate Prediction Center Morphing (CMORPH) and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) to generate a high resolution and advanced satellite product, called, Integrated Multi-satellitE Retrievals for the Global Precipitation Mission (IMERG) [8,30].

The Global Precipitation Measurement (GPM) mission is an international network of satellites using a core satellite combined with ancillary ones and an advanced passive microwave radiometer paired with dual-frequency radar system to measure rain and snow from space. Data are collected from a constellation of satellites such as MetOp B/C (operated by European Organization for the Exploitation of Meteorological Satellites (EUMESAT)), Global Change Observation Mission-Water (GCOM-W) (operated by Japan Aerospace Exploration Agency (JAXA)), Megha-Tropiques (mutually operated by Indian Space Research Organization (ISRO) and French Centre National d'Etudes Spatiales (CNES) and some other satellites, mainly managed by NASA [31,32]. Similar to previous evaluation studies of GPM predecessors such as CMORPH [33,34], PERSIANN [35,36], GSMaP [37,38], CHIRPS [39], and TRMM [40,41], some efforts have been made to evaluate and to characterize uncertainties and errors of this satellite-based product [12,42–45]. A significant improvement of GPM over other satellite-based products is its ability to detect light and solid precipitation by combining the GPM Microwave Imager (GMI) and Dual-Frequency Precipitation Radar (DPR) instruments [12].

Some of the previous verification studies for IMERG products were mainly conducted over very large regions in China e.g., [15] or India e.g., [30]. Consequently, China Gauge-based Daily Precipitation Analysis (CGDPA) ($0.25^\circ \times 0.25^\circ$), hourly gridded precipitation data from the China Meteorological Administration (CMA) ($0.1^\circ \times 0.1^\circ$), and the gridded Indian Meteorological Department (IMD) gauge-based rainfall data have been used as trustable sources (benchmark or ground truth) of observational precipitation records. These gauge-based products undergo a strict quality control process before publishing [5,15]. However, by comparing them to IMERG products can cause some computing issues. Since the Inverse Distance Weighting (IDW) method is generally used to interpolate precipitation values among gauge observations to produce gridded data for large areas, it can cause uncertainty in precipitation records estimates over regions with sparse gauge networks [15]. Moreover, spatiotemporal resolutions for these products are generally coarser than IMERG spatiotemporal resolutions. As a result, some data will be missed while matching products with finer resolution with products with relatively coarser resolution [46]. Therefore, applying a more robust gridded source of precipitation (higher spatiotemporal resolutions) can lead to a better assessment of IMERG product comparing to previous studies.

Moreover, to perform a comprehensive validation, there is a need to evaluate the IMERG satellite product performance at regional and local scales and under adverse weather conditions such as hurricanes or thunderstorms. During May 2015, natural disasters in the U.S. caused more than $1 billion in damages, while Texas was among the most vulnerable States [12,47]. In August 2017, Hurricane Harvey, as a category 4 major hurricane event, inflicted more than $125 billion in damages to Texas in a week [48,49]. Texas will likely experience similar events during future years. Hence, it is critical to have a powerful tracing mechanism to detect events and their characteristics

from space. The objective of this study was to evaluate IMERG product at a regional scale in an area which is extremely vulnerable to flood and hurricane events and show how the GPM core satellite can capture hurricane events. It should be noted that other two GPM satellite products (Early- and late-run products) can also be utilized in near real-time applications for forecasting purposes. To fulfill this goal, weather radar data as observational data was compared with an open-source satellite-based precipitation product. Section 2 provides the study area, utilized datasets and methodology. In Section 3, results are provided and comprehensively discussed while Section 4 represents the conclusion.

2. Materials and Methods

2.1. Study Area

Texas, the largest state in the contiguous United States of America (268,597 mi^2), is one of the most flood-prone and hurricane affected regions in the world [50]. It is generally divided into three major climatic zones with different hydrometeorological conditions, topographic/physiographic characteristics, and land use/land cover (LULC) types: (1) wet (eastern region) (2) semi-wet (central region) (3) dry (western region) [51]. The western region is typically covered by arid deserts while the central and eastern districts are generally more humid covered by forest, agricultural lands, and rangelands. The southeast is dominantly under the effect of subtropical weather condition which causes this region to be renowned as one of the most vulnerable regions to natural disasters such as hurricanes, tornados and flash-floods [52]. The majority of Texas population resides in the central and eastern regions causing growth demand for more infrastructures in this area. Additionally, due to the natural disasters, Texas suffers a significant financial loss every year across these regions. Consequently, the study area covers from 93.4°–106.7° West in longitude and 25.7°–36.6° North in latitude.

2.2. Precipitation Data

2.2.1. NWS/NCEP Stage-IV Radar Data

The National Weather Service/National Centers for Environmental Prediction (NWS/NCEP) River Forecast Center (RFC) Stage-IV Quantitative Precipitation Estimates (QPEs) gridded radar-rainfall (http://data.eol.ucar.edu/) was employed as the source of precipitation observation data. The operational stage IV precipitation radar data production began in 2001 and has been continued since then [53]. It contains adjusted hourly rain-gauge precipitation records with a 4 × 4-km spatial resolution and hourly temporal resolution. The measurements are in cumulative hourly format at the end of each hour.

2.2.2. IMERG Satellite Product

The level 3 IMERG data was downloaded from the Precipitation Measurement Missions website (http://pmm.nasa.gov/data-access/downloads/gpm). Version 05-B of IMERG product was used with the half-hourly temporal resolution and 0.1° × 0.1°(11 × 11 km) spatial resolution. The final-run product is adjusted with gauge-network which converts it to more accurate product relative to two other early and late-run products. It is produced almost 3 months after the date of preliminary records and mainly used for research purposes. Since the radar products have an hourly temporal resolution, the IMERG products used in this study were also converted to hourly format at the end of each hour.

2.3. Methodology

Hurricane Harvey predominantly affected the southern and eastern counties in the state of Texas. Therefore, to better track hurricane Harvey and accurately detect the spatial extent of the event; only those grids with more than 10 mm cumulative precipitation captured by radar product were evaluated over the entire state during the event. The gridded final-run IMERG product was compared

to the nearest NCEP Stage IV gridded radar product at hourly temporal resolution for five consecutive days during hurricane Harvey. The nearest radar grid to each satellite product was identified based on the Euclidean distance and paired to each satellite grid. Since the spatial resolution of satellite product is almost three times the radar grids, for each satellite product, there is a unique nearest radar grid. Afterward, the precipitation records from the IMERG grids were compared with that corresponding to the nearest radar grids at hourly temporal resolution to calculate the statistical comparative metrics. There are 45, 509 NWS/NCEP radar grids versus 6502 IMERG satellite grids inside Texas. However, just 2508 grids are included in the region with cumulative precipitation magnitude of over 10 mm.

Cumulative rainfall rates show discrepancies in spatial patterns between the stage IV gridded product and the IMERG satellite product for the same grids during the study period (Figure 1). IMERG reveals the total precipitation approximately twice the NCEP stage-IV radar over the southeast of the state.

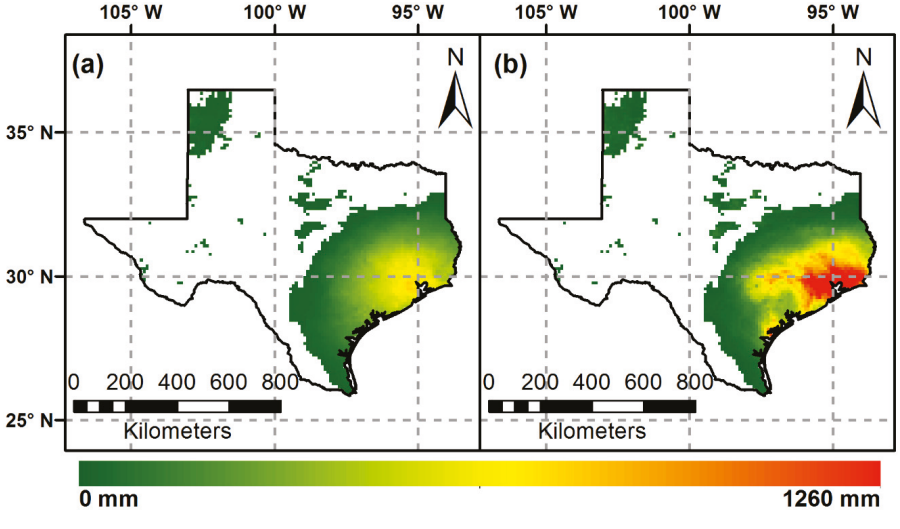


Figure 1. Cumulative precipitation measurement (mm) and spatial pattern captured by (**a**) NCEP stage-IV Radar (**b**) GPM-IMERG satellite final-run product, during Hurricane Harvey from 25 to 29, August 2018.

Index Classification

Two sets of performance measures were used for assessment of the IMERG product (Table 1). The first set includes Pearson Correlation Coefficient (CC), Mean Error (ME), Relative Bias (BIAS), Mean Bias Factor (MBF), and Root Mean Squared Error (RMSE). These statistical criteria mainly investigate the consistency of data based on reference precipitation radar data. The second set of indices provides information about the probability and accuracy of precipitation detection by IMERG product compared to reference radar products. Probability Of (rainfall) Detection (POD), False Alarm Ration (FAR), Critical Success Index (CSI), and Peirce Skill Score (PSS) are widely used to assess the probabilistic quality of satellite products in this category [12,30]. POD indicates the probability of rainfall detection by satellite while reference radar data reports the precipitation. FAR index explains the inaccuracy of the satellite falsely detecting rainfall, when radar has not detected the rainfall. CSI combines all events captured by both satellite and radar, including the satellite's missed and false events according to reference radar observations. It shows more harmony between radar and satellite rainfall detection, while providing more deterministic vision of satellite data based on various spatiotemporal resolutions [12]. PSS is a difference between probability of detection and probability of

false detection and indicates the system's ability to capture rainfall values. The perfect value is 1 for this index while it can range between −1 and 1 [54,55].

Table 1. Statistical Performance Measures used in evaluation and comparison.

Index	Formula	Range	Perfect Value
Category 1 [1] (Basic Statistical Indices)			
Correlation Coefficient (CC) [2]	$\frac{\frac{1}{N}\sum_{n=1}^{N}(Sat_n-\overline{Sat})(Rad_n-\overline{Rad})}{(SD_{Sat})(SD_{Rad})}$	(−1)–(+1)	1
Mean Error (ME)	$\frac{1}{N}\sum_{n=1}^{N}(Sat_n - Rad_n)$	$-\infty/+\infty$	0
Relative Bias (RBIAS)	$\frac{\sum_{n=1}^{N}(Sat_n-Rad_n)}{\sum_{n=1}^{N}Rad_n}$	$-\infty/+\infty$	0
Mean Bias Factor (MBF)	$\frac{\sum_{n=1}^{N}Rad_n}{\sum_{n=1}^{N}Sat_n}$	0–+∞	1
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(Sat_n - Rad_n)^2}$	0–+∞	0
Category 2 [1] (Probabilistic Statistical Indices)			
Probability of detection (POD)	$\frac{C_{SR}}{C_{SR}+C_RM_S}$	0–1	1
False Alarm Ration (FAR)	$\frac{C_SM_R}{C_{SR}+C_SM_R}$	0–1	0
Critical Success Index (CSI)	$\frac{C_{SR}}{C_{SR}+C_SM_R+C_RM_S}$	0–1	1
Peirce Skill Score	$\frac{C_{SR}}{C_{SR}+C_RM_S}-\frac{C_SM_R}{M_{SR}+C_SM_R}$	(−1)–(+1)	1

[1] n: number of samples; Sat_n: satellite rainfall estimate; Rad_n: radar-observed rainfall; $\overline{Sat}$: averaged satellite rainfall; $\overline{Rad}$: averaged radar rainfall; SD_{Sat}: standard deviations of satellite rainfall; SD_{Rad}: standard deviations of radar rainfall; C_{SR}: events captured by satellite and radar; C_SM_R events captured by satellite and missed by radar; C_RM_S: events captured by radar and missed by satellite; M_{SR}: events missed by radar and satellite (no event). [2] The CC usually ranges from 0 to 1. However, to better show the discrepancies, the negative correlation coefficients were utilized in this study.

3. Results and Discussion

The cumulative precipitation spatial pattern for both products provides more insight on how satellite and radar products can capture extreme rainfall (Figure 1). The cumulative precipitation rate captured by radar product ranged between 10 and 770 mm; while, the IMERG product ranged between 0 and 1260 mm for the same regions (Figure 1). Comparison of the two data products shows that the satellite product generally overestimated the precipitation rate while it failed to detect precipitation records in some regions. Both products detected the hurricane center at the east side of Texas, near Harris County (Houston Metropolitan), while their spatial patterns showed some discrepancies. In general, while radar products are more concentrated over the hurricane core; satellite products have broader extent and detected more southern cities (Rockport, Port Aransas and Corpus Christi) affected by the hurricane. Based on local observations and reports, satellite products may be able to trace the hurricane's spatial pattern and path more reliably than radar products. Moreover, it showed that satellite products have generally better performance over grids with higher precipitation rates compared to grids with lower rates.

3.1. Basic Statistical Indices

The CC spatial pattern between radar and satellite products during hurricane Harvey (25–29 August 2018) provides more details of satellite product consistency (Figure 2). Due to the higher spatial variability rate during hurricanes, tracking the correlation coefficient can reveal the performance of satellite products and illustrates hurricane characteristics.

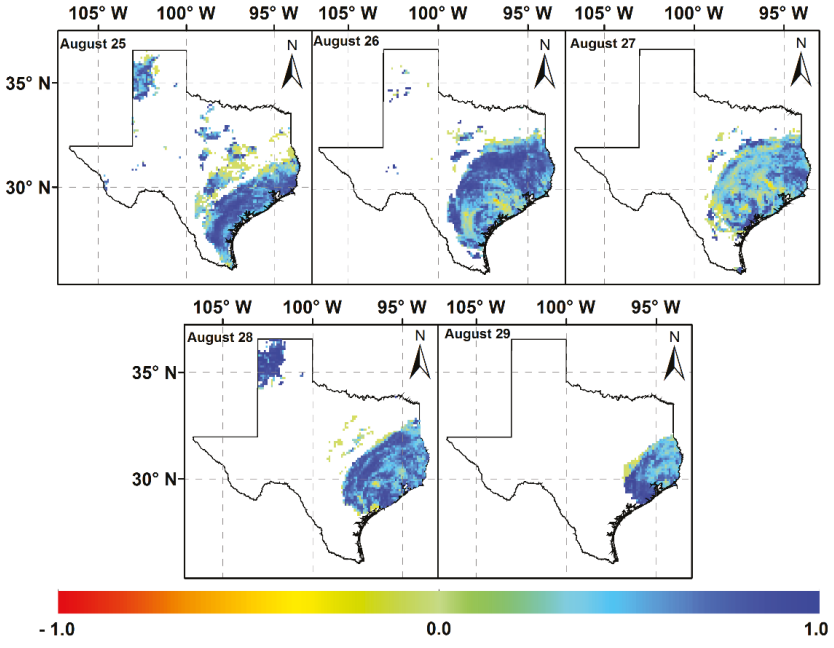


Figure 2. Spatial distributions of the Correlation Coefficient (CC) factor for hourly IMERG precipitation product at 0.1° × 0.1° resolution over Texas from 25 to 29 August 2017.

Satellite products had lower consistency with the radar products near the center of hurricane; while, moving away from the center, the consistency improved (Figure 2). Additionally, it illustrates that although number of grids with the perfect CC value (1) is higher during day two, a higher level of inconsistency with the lowest recorded CC values for satellite product (−0.57) also occurred. It can reveal that a complex storm structure may cause inconsistency in measuring precipitation rates by IMERG product even during the same day and nearby cells. Two important issues should be noted: (1) the perfect value for CC was obtained for the first two days with higher rates of precipitation compared to the three days in which the precipitation rate decreased around 50% and (2) the hurricane complex structure can substantially affect the satellite performance especially at the hurricane core.

The scatterplot between IMERG satellite product and NCEP stage-IV radar product during hurricane Harvey provides more details of consistency between the two products during the event (Figure 3). It represents the correlation among the observations and estimations for each day during the hurricane span. Higher intensity rates during 26 and 27 August caused more expansion in scatter plots and lower correlation rates among the satellite data and the observations. Hence, the scatterplot provides a better vision on how the hurricane advanced to the category IV storm during this span and subsequently was recognized as one of the most destructive hurricanes in the United States. On 26 August, the day that the event advanced to the category IV hurricane with the maximum intensity and duration, the satellite product demonstrated the lowest correlation to the radar observations (Figure 3). It shows that satellite product has lower efficiency while capturing an excessive amount of rainfall due to the hurricane. At the end of the 5-day span (29 August), the correlation slightly improved. It reveals that IMERG product algorithm still needs improvement to fully capture the extreme rainfall events.

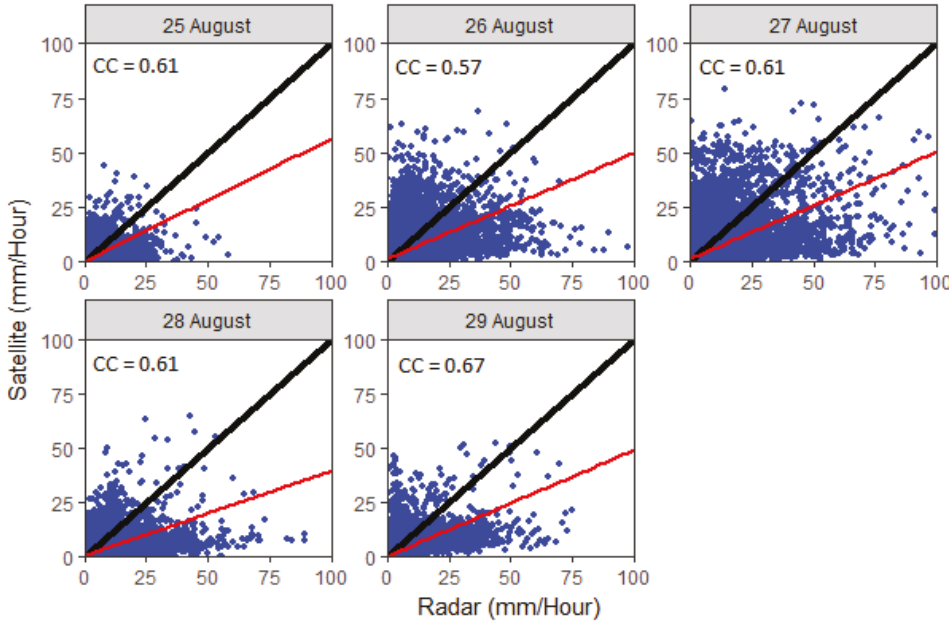


Figure 3. Scatterplots of IMERG satellite product and stage-IV NCEP radar precipitation product with an hourly temporal resolution between 25th and 29th August 2017.

Box plots of probabilistic (POD, FAR, CSI and PSS) and basic (ME, MBF, BIAS and RMSE) statistical indices over grid boxes of the IMERG hourly precipitation product provide more details of satellite data distribution (Figure 4). The upper and lower edges of the central box are the first (25%) and third (75%) quantiles that represent the interquartile range (IQR), while the central inside band and cross sign represents the median and mean values respectively. The minimum and maximum observations which are usually selected among 10 to 90 percent of measurements are 1.5x(IQR) from the upper and lower quartiles. Farther observations show the outliers which can reveal inconsistency level between satellite and radar observations. Basic indices have strong symmetry toward the median band while probabilistic indices show wider distribution ranges. However, basic indices have more outliers compared to probabilistic ones. Additionally, basic indices have generally upper-band outliers (except ME), while probabilistic indices are commonly associated with lower-band outliers (Figure 4). It explains that satellite product overestimated precipitation values while they commonly underestimated event's frequency and missed events. There are no outliers for FAR and CSI, while the POD and PSS have only lower outliers. The results show that although IMERG product may falsely report non-existing events or misses true events, its general performance was acceptable in detecting precipitation events as the POD's upper quantile, median and mean values are very close to their perfect thresholds. In the case of basic indices, the distributions are wider but more symmetric. The distance between upper and lower quantiles are smaller while the outliers show that, during heavy events like hurricanes, the performance of IMERG can be weaker compared to more frequent events. While ME shows two-way outliers, MBF, BIAS and RMSE just have upper outliers and generally shows overestimation of satellite products. The general view to the box plots shows that during hurricanes IMERG generates more truncated data and dominantly overestimates the measurements.

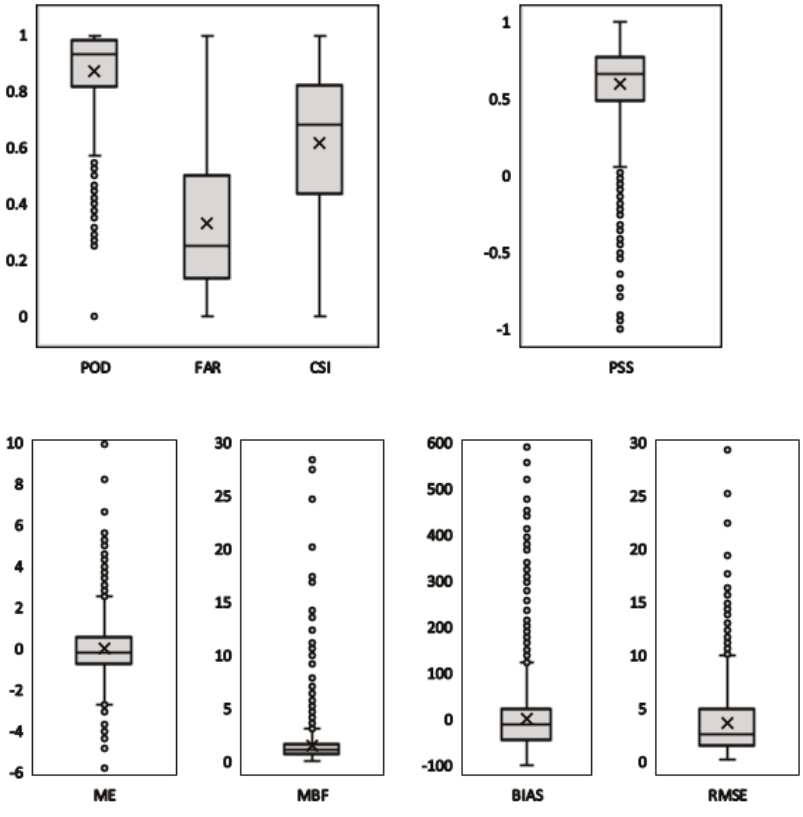


Figure 4. Box plots of probabilistic (POD, FAR, CSI and PSS) and basic (ME, MBF, BIAS and RMSE) statistical indices over grid boxes for IMERG hourly precipitation products.

ME, BIAS, MBF and RMSE spatial variability and their values illustrates more details for IMERG product during hurricane Harvey (Figure 5).

The Mean Error (ME) index provides the estimation performance of the satellite product over the affected area by the hurricane (Figure 5a). It generally reveals that the satellite product overestimated precipitation over the regions with higher precipitation rates (mainly near Gulf of Mexico). For the regions with lower precipitation rates, the ME was mostly near its perfect value (0) or had negative values which shows some underestimation by the satellite product. Based on hurricane Harvey spatial variability (Figure 1), ME spatial variability shows that IMERG final-run product generally follows the same pattern. Moreover, although IMERG indicated over- and under-estimations during hurricane Harvey, the range (−6 to 10 mm) is within a reasonable range, which indicates the good performance of the IMERG product during the hurricane event. Figure 5b shows spatial variability of BIAS over the area affected by hurricane Harvey. This index also follows the same pattern as ME index, while it shows very good performance over regions with supercritical precipitation. The BIAS index was near zero or had positive values in more affected areas by hurricane, while other regions mainly had negative values. The mean bias factor (MBF) generally shows the good performance of IMERG product based on radar observations (Figure 5c). It shows that the IMERG product had generally reasonable estimation of precipitation rate. Figure 5d provides spatial distribution of root mean square error (RMSE) during the event. It presents the higher RMSE error rates over more eastern part with supercritical precipitation and near the Mexican Gulf. RMSE spatial pattern also strengthens this

requirement/finding that IMERG final-run product algorithm needs improvement to better estimate precipitation rate during hurricanes.

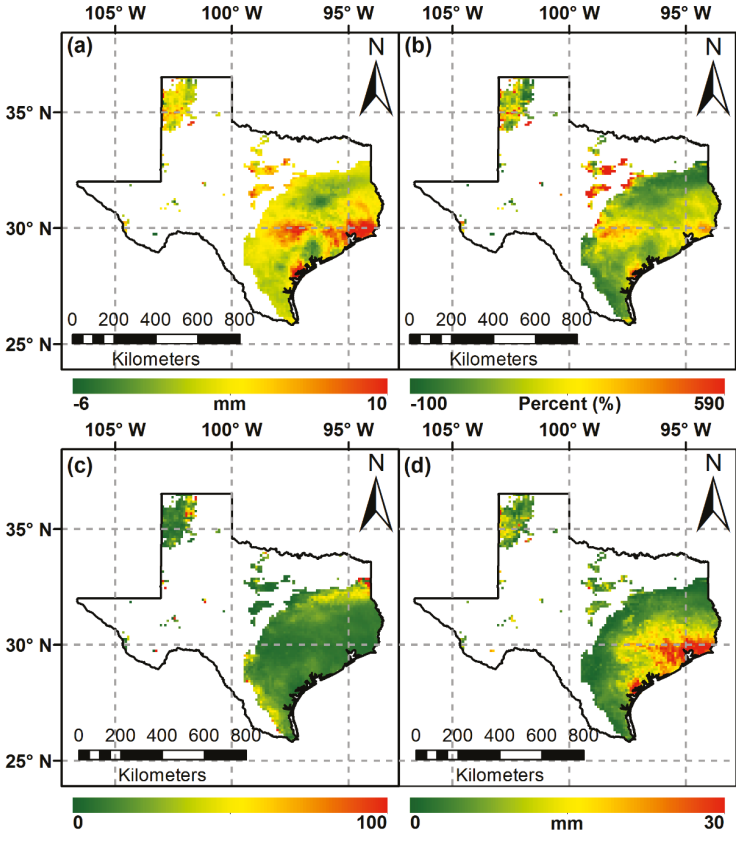


Figure 5. Spatial distribution of basic statistical indices for the final-run IMERG hourly product during Hurricane Harvey (**a**) ME (**b**) BIAS (**c**) MBF (**d**) RSME.

3.2. Probabilistic Statistical Indices

POD, CSI, FAR and PSS spatial patterns and their values for IMERG product during hurricane Harvey provide more details of probabilistic characteristics of satellite products (Figure 6).

IMERG products can satisfactorily detect precipitation in most parts hit by the event (Figure 6a). Most of the grids showed (nearly) perfect POD values. However, between southern and central regions (between Austin and San Antonio), the probability of detection rate was less than the others which may be due to the lower rate of precipitation in these areas and the weaker capability of IMERG in detecting lower precipitation rates. Moreover, POD values beyond the border line (grids with less than 10 mm precipitation during the hurricane event) are significantly lower, which shows that satellite products cannot correctly detect the precipitation events with lower intensities during hurricane events. Figure 6b shows the CSI values. It almost followed the same spatial pattern as POD with the smoother values. CSI values (contrary to POD values) are not close to their perfect amount in southern districts. POD and CSI values demonstrate that IMERG product can significantly capture hurricane events at the regions with higher intensities comparing to regions farther from the hurricane core. FAR values over the hurricane affected areas are near their perfect value, 0 (Figure 6c). It shows that in other

regions most of the hits detected by satellite were also captured by radar. It shows that the IMERG product has the general ability to correctly detect precipitation hits during hurricanes. PSS values are shown in Figure 6d. It generally shows that the IMERG product has a reasonable performance in estimating rainfall. However, negative PSS values in the regions with the higher rate of precipitation (eastern parts-hurricane eye) illustrates the relative weakness of IMERG algorithm to capture extreme precipitation rates. Additionally, farther from the hurricane center, the PSS values are closer to their perfect value. It can lead to the conclusion that IMERG ability in estimating rainfall over the majority of affected areas by the hurricane is reasonable.

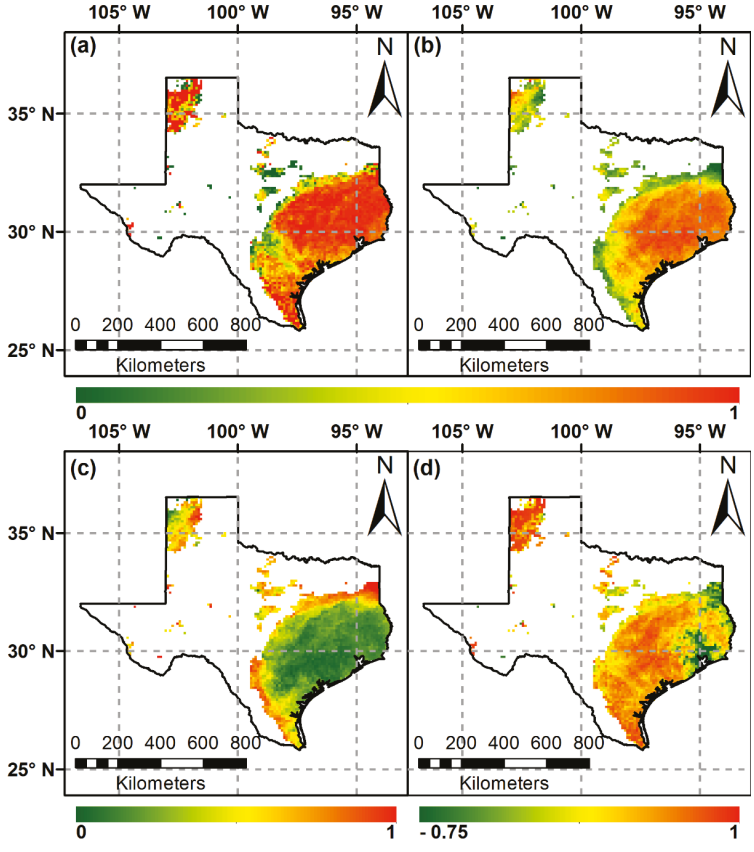


Figure 6. Spatial distribution of probabilistic statistical indices for the final-run IMERG during hurricane Harvey (**a**) POD (**b**) CSI (**c**) FAR (**d**) PSS.

4. Conclusions

In this study, the performance of GPM-IMERG satellite final-run product regarding precipitation monitoring during hurricane events was analyzed. Hurricane Harvey, recognized as a Category 4 storm by National Hurricane Center (NHC), National Oceanic and Atmospheric Administration (NOAA), occurred during 25–29 August 2017. The quality of IMERG product at hourly and $0.1^\circ \times 0.1^\circ$ resolutions to capture precipitation during the hurricane event was evaluated. For this purpose, hourly precipitation observation data from NCEP stage IV radar product adjusted by ground gauges, was used. To better capture the hurricane event, the analysis was only done over radar grids with cumulative rainfall records of over 10 mm. Due to the high rate of hurricane spatial variability, the data

was comprehensively analyzed during five consecutive days. The hurricane increased to category 4 during 26 and 27 August. The results suggest that IMERG can appropriately detect and trace hurricane spatial path and variability while a better performance for the precipitation rates is needed. The main conclusions are given as follows:

1. The general evaluation demonstrates that IMERG product can accurately detect and trace hurricane spatial pattern while the estimation algorithm needs to be improved to better measure the precipitation intensity.
2. The IMERG hourly precipitation product shows significant overestimation over the storm's peak regions dominantly near to the coast. This overestimation gradually decreases away from the hurricane center. The basic statistical indices generally reflect this overestimation of the satellite product, however, the small bias ($\pm$10%) over regions with the precipitation peak can smooth the unsatisfactory performance of satellite products based on these metrics.
3. Statistical indices demonstrate an adequate performance of satellite products in detection of precipitation over the area affected by hurricane. Most of the area shows high POD (>0.8) value associated with low FAR (<0.2) which validates the satellite performance regarding the predictability of rainfall hits and not reporting false hits.
4. CC spatial distributions during five consecutive days reveal that when the hurricane advanced to the category 4 storm (2nd day/26 August); although most of the sub-regions showed high CC values, near the center of the hurricane (hurricane eye), there is negative correlation. It indicates the complex internal structure and spatial variability of the storm was not well captured by the satellite. Additionally, lower quality input data from multiple sensors can intensify the inconsistency of the satellite products. Therefore, deeper understanding of IMERG product's diurnal cycle may help to generate better algorithm for estimating precipitation records in future.
5. The CSI and PSS indices generally reflected a satisfactory performance by the satellite products, however, for the sub-regions especially near to the eastern coast, IMERG could not capture the storm appropriately.

Based on the results of this study, the quality of IMERG product can be improved to estimate precipitation during extreme events such as hurricane or over sub-regions with extreme intensity rainfall. The results of the detailed analysis in this study can be utilized by atmospheric and earth scientists to better evaluate satellite products' performance during hurricane events and improve forecasting algorithms. Accurate estimates of hurricanes (Category IV/V) can undoubtedly help to increase society's safety and plan for better optimized evacuation of residents from areas affected by hazardous storms.

Author Contributions: E.O. and H.O.S. conceived and designed the methodology; E.O. analyzed the data; A.A.T. provided reviews and suggestions; E.O. prepared and wrote the paper with the multiple inputs and reviews by H.O.S. and A.A.T.

Funding: This research was partially funded by U.S. Army Research Office (Grant W912HZ-14-P-0160) and a grant from UTSA (University of Texas at San Antonio) Open Cloud Institute (OCI) entitled "Prediction of Crash Incidence and Severity on Texas Roadways".

Acknowledgments: These supports are cordially acknowledged.

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

References

1. Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]
2. 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.* **2011**, *13*, 67–83. [CrossRef]
3. 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]

4. 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]
5. Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Liu, Z.; Norouzi, H.; Pai, D.S. A preliminary assessment of gpm-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.* **2018**, *556*, 865–876. [CrossRef]
6. Afshari, S.; Tavakoly, A.A.; Rajib, M.A.; Zheng, X.; Follum, M.L.; Omranian, E.; Fekete, B.M. Comparison of new generation low-complexity flood inundation mapping tools with a hydrodynamic model. *J. Hydrol.* **2018**, *556*, 539–556. [CrossRef]
7. Lane, J.; Kasparis, T.; Michaelides, S.; Metzger, P. A phenomenological relationship between vertical air motion and disdrometer derived a-b coefficients. *Atmos. Res.* **2018**, *208*, 94–105. [CrossRef]
8. 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]
9. Liu, G.; Schwartz, F.W. Climate-driven variability in lake and wetland distribution across the prairie pothole region: From modern observations to long-term reconstructions with space-for-time substitution. *Water Resour. Res.* **2012**, *48*. [CrossRef]
10. Javaheri, A.; Nabatian, M.; Omranian, E.; Babbar-Sebens, M.; Noh, J.S. Merging real-time channel sensor networks with continental-scale hydrologic models: A data assimilation approach for improving accuracy in flood depth predictions. *Hydrology* **2018**, *5*. [CrossRef]
11. 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]
12. 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**. [CrossRef]
13. Afshari, S.; Omranian, E.; Feng, D. Relative sensitivity of flood inundation extent by different physical and semi-empirical models. *Natl. Water Center Innov. Program Summer Inst. Tech. Rep.* **2016**, 19–24. [CrossRef]
14. Xie, P.; Arkin, P.A. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Clim.* **1996**, *9*, 840–858. [CrossRef]
15. 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]
16. 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]
17. Mishra, A.K.; Coulibaly, P. Developments in hydrometric network design: A review. *Rev. Geophys.* **2009**, *47*. [CrossRef]
18. Germann, U.; Galli, G.; Boscacci, M.; Bolliger, M. Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.* **2006**, *132*, 1669–1692. [CrossRef]
19. Heinselman, P.L.; Priegnitz, D.L.; Manross, K.L.; Smith, T.M.; Adams, R.W. Rapid sampling of severe storms by the national weather radar testbed phased array radar. *Weather Forecast.* **2008**, *23*, 808–824. [CrossRef]
20. Aksoy, A.; Dowell, D.C.; Snyder, C. A multicase comparative assessment of the ensemble kalman filter for assimilation of radar observations. Part i: Storm-scale analyses. *Mon. Weather Rev.* **2009**, *137*, 1805–1824. [CrossRef]
21. Gabella, M.; Morin, E.; Notarpietro, R.; Michaelides, S. Winter precipitation fields in the southeastern mediterranean area as seen by the ku-band spaceborne weather radar and two c-band ground-based radars. *Atmos. Res.* **2013**, *119*, 120–130. [CrossRef]
22. Sharif, H.O.; Ogden, F.L.; Krajewski, W.F.; Xue, M. Numerical simulations of radar rainfall error propagation. *Water Resour. Res.* **2002**, *38*, 15-11–15-14. [CrossRef]
23. Mei, Y.; Anagnostou, E.N.; Nikolopoulos, E.I.; Borga, M. Error analysis of satellite precipitation products in mountainous basins. *J. Hydrometeorol.* **2014**, *15*, 1778–1793. [CrossRef]
24. Sharif, H.O.; Ogden, F.L. Mass-conserving remapping of radar data onto two-dimensional cartesian coordinates for hydrologic applications. *J. Hydrometeorol.* **2014**, *15*, 2190–2202. [CrossRef]
25. 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]

26. Hossain, F.; Katiyar, N. Advancing the use of satellite rainfall datasets for flood prediction in ungauged basins: The role of scale, hydrologic process controls and the global precipitation measurement mission. In *Quantitative Information Fusion for Hydrological Sciences*; Cai, X., Yeh, T.C.J., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 79, pp. 163–181.
27. Hobouchian, M.P.; Salio, P.; García Skabar, Y.; Vila, D.; Garreaud, R. Assessment of satellite precipitation estimates over the slopes of the subtropical andes. *Atmos. Res.* **2017**, *190*, 43–54. [CrossRef]
28. Salio, P.; Hobouchian, M.P.; García Skabar, Y.; Vila, D. Evaluation of high-resolution satellite precipitation estimates over southern south america using a dense rain gauge network. *Atmos. Res.* **2015**, *163*, 146–161. [CrossRef]
29. Retalis, A.; Tymvios, F.; Katsanos, D.; Michaelides, S. Downscaling chirps precipitation data: An artificial neural network modelling approach. *Int. J. Remote Sens.* **2017**, *38*, 3943–3959. [CrossRef]
30. 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]
31. Karbalaee, N.; Hsu, K.; Sorooshian, S.; Braithwaite, D. Bias adjustment of infrared-based rainfall estimation using passive microwave satellite rainfall data. *J. Geophys. Res.-Atmos.* **2017**, *122*, 3859–3876. [CrossRef]
32. 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]
33. Gaona, M.F.R.; Overeem, A.; Leijnse, H.; Uijlenhoet, R. First-year evaluation of gpm rainfall over the netherlands: Imerg day 1 final run (v03d). *J. Hydrometeorol.* **2016**, *17*, 2799–2814. [CrossRef]
34. 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]
35. 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]
36. 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]
37. Habib, E.; Haile, A.T.; Tian, Y.; Joyce, R.J. Evaluation of the high-resolution cmorph satellite rainfall product using dense rain gauge observations and radar-based estimates. *J. Hydrometeorol.* **2012**, *13*, 1784–1798. [CrossRef]
38. Chen, S.; Liu, H.; You, Y.; Mullens, E.; Hu, J.; Yuan, Y.; Huang, M.; He, L.; Luo, Y.; Zeng, X.; et al. Evaluation of high-resolution precipitation estimates from satellites during july 2012 beijing flood event using dense rain gauge observations. *PLoS ONE* **2014**. [CrossRef] [PubMed]
39. Miao, C.; Ashouri, H.; Hsu, K.; Sorooshian, S.; Duan, Q. Evaluation of the persiann-cdr daily rainfall estimates in capturing the behavior of extreme precipitation events over china. *J. Hydrometeorol.* **2015**, *16*, 1387–1396. [CrossRef]
40. Katiraie-Boroujerdy, P.-S.; Nasrollahi, N.; Hsu, K.; Sorooshian, S. Evaluation of satellite-based precipitation estimation over iran. *J. Arid Environ.* **2013**, *97*, 205–219. [CrossRef]
41. 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*. [CrossRef]
42. Tian, Y.; Peters-Lidard, C.D.; Adler, R.F.; Kubota, T.; Ushio, T. Evaluation of gsmap precipitation estimates over the contiguous united states. *J. Hydrometeorol.* **2009**, *11*, 566–574. [CrossRef]
43. Katsanos, D.; Retalis, A.; Tymvios, F.; michaelides, S. Analysis of precipitation extremes based on satellite (chirps) and in situ dataset over cyprus. *Nat. Hazards* **2016**, *83*, 53–63. [CrossRef]
44. Chen, S.; Hong, Y.; Cao, Q.; Gourley, J.J.; Kirstetter, P.; 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*, 13–060. [CrossRef]
45. 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]
46. Shrestha, R.; Houser, P.R.; Anantharaj, V.G. An optimal merging technique for high-resolution precipitation products. *J. Adv. Model. Earth Syst.* **2011**, *3*. [CrossRef]

47. Wolter, K.; Eischeid, J.K.; Cheng, L.; Hoerling, M. What history tells us about 2015 US daily rainfall extremes. *Bull. Am. Meteorol. Soc.* **2016**, *97*, S9–S13. [CrossRef]
48. Craig, R.K. Harvey, Irma, and the NFIP: Did the 2017 Hurricane Season Matter to Flood Insurance Reauthorization? Available online: https://ssrn.com/abstract=3129088 (accessed on 23 February 2018).
49. Costliest U.S. Tropical Cyclones Tables Updated. Available online: https://www.nhc.noaa.gov/news/UpdatedCostliest.pdf (accessed on 26 January 2018).
50. Omranian, E.; Sharif, H.; Dessouky, S.; Weissmann, J. Exploring rainfall impacts on the crash risk on texas roadways: A crash-based matched-pairs analysis approach. *Accid. Anal. Prev.* **2018**, *117*, 10–20. [CrossRef] [PubMed]
51. Nielsen-Gammon, J.W.; Zhang, F.; Odins, A.M.; Myoung, B. Extreme rainfall in texas: Patterns and predictability. *Phys. Geogr.* **2005**, *26*, 340–364. [CrossRef]
52. Furl, C.; Sharif, H.; Zeitler, J.W.; Hassan, A.E.; Joseph, J. Hydrometeorology of the catastrophic Blanco river flood in South Texas, May 2015. *J. Hydrol. Reg. Stud.* **2018**, *15*, 90–104. [CrossRef]
53. Lin, Y.; Mitchell, K.E. The NCEP Stage II/IV Hourly Precipitation Analysis: Development and Applications. Available online: https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm (accessed on 12 June 2018).
54. Dos Reis, B.J.; Rennó, D.C.; Lopes, S.E. Validation of satellite rainfall products over a mountainous watershed in a humid subtropical climate region of brazil. *Remote Sens.* **2017**, *9*. [CrossRef]
55. Manzato, A. A note on the maximum peirce skill score. *Weather Forecast.* **2007**, *22*, 1148–1154. [CrossRef]

remote sensing

MDPI

Article

Developing an Ensemble Precipitation Algorithm from Satellite Products and Its Topographical and Seasonal Evaluations Over Pakistan

Khalil Ur Rahman, Songhao Shang *, Muhammad Shahid and Jiang Li

State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China; hlrhm17@mails.tsinghua.edu.cn (K.U.R.); shahidbaloch21@gmail.com (M.S.); lj397947820@163.com (J.L.)

* Correspondence: shangsh@tsinghua.edu.cn; Tel.: +86-10-6279-6674

Received: 15 October 2018; Accepted: 14 November 2018; Published: 20 November 2018

Abstract: Accurate estimation of precipitation is critical for hydrological, meteorological, and climate models. This study evaluates the performance of satellite-based precipitation products (SPPs) including Global Precipitation Measurement (GPM)-based Integrated Multi-Satellite Retrievals for GPM (IMERG), Tropical Rainfall Measurement Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA 3B43-v7), Precipitation Estimation from Remotely-Sensed Information using Artificial Neural Network (PERSIANN), and PERSIANN-CDR (Climate Data Record), over Pakistan based on Surface Precipitation Gauges (SPGs) at spatial and temporal scales. A novel ensemble precipitation (EP) algorithm is developed by selecting the two best SPPs using the Paired Sample *t*-test and Principal Component Analysis (PCA). The SPPs and EP algorithm are evaluated over five climate zones (ranging from glacial Zone-A to hyper-arid Zone-E) based on six statistical metrics. The result indicated that IMERG outperformed all other SPPs, but still has considerable overestimation in the highly elevated zones (+20.93 mm/month in Zone-A) and relatively small underestimation in the arid zone (−2.85 mm/month in Zone-E). Based on the seasonal evaluation, IMERG and TMPA overestimated precipitation during pre-monsoon and monsoon seasons while underestimating precipitation during the post-monsoon and winter seasons. However, the developed EP algorithm significantly reduced the errors both on spatial and temporal scales. The only limitation of the EP algorithm is relatively poor performance at high elevation as compared to low elevations.

Keywords: satellite precipitation; Global Precipitation Measurement (GPM); IMERG; TRMM-TMPA; Ensemble Precipitation (EP) algorithm; topographical and seasonal evaluation

1. Introduction

Precipitation is crucial input parameter of the global hydrological cycle [1,2] and an impetuous factor contributing to natural disasters like droughts and flooding [3]. The performance of different hydrological, meteorological, and climate models depends on the accuracy of precipitation inputs. These models are used in reliable modeling, monitoring, and quantification of floods, drought assessment, landslides, agricultural production, and sustainable water resource management. Moreover, understanding of the spatial and temporal variability of precipitation is significantly important under climate change [4,5]. The spatial and temporal variation in precipitation pattern significantly affects socioeconomic factors such as disaster management, food security, ecosystem health, and hydropower generation [6–8]. Therefore, accurate precipitation estimation with high spatiotemporal resolution on a regional scale is essential for significant hydrological predictions. However, this is still a challenging task for the developing countries like Pakistan because of the sparse surface precipitation gauge (SPG) network and highly complex topography [9–11].

Most widely used techniques for precipitation measurement are in situ measurements (SPGs), and commercial microwave products such as ground-based radars and satellite-based sensors [4,12]. There are numerous errors associated with distribution and readings of SPG as discussed in [13]. Furthermore, the SPG measurements are in situ, which is difficult to describe the spatial variations in precipitation on a regional scale. To overcome these uncertainties, the utilization of satellite-based precipitation products (SPPs) to measure global or regional precipitation has increased significantly over the past thirty years [14]. In poorly or ungauged regions, SPP estimation techniques may be opted due to unavailability, uncertainty, and quantitative self-consistency of SPG data [15]. At present, the development of high-resolution SPPs has provided unprecedented opportunity to monitor the spatiotemporal variability in precipitation on a global scale, particularly at high elevation where ground-based information is scarce or not readily available [16–18]. A detailed description of the most widely used SPPs can be found in a past work [19].

A number of researchers evaluated the performance of these SPPs on the regional and global scales, such as Asia [3,20–23], Africa [24–26], North and South America [27–30], Europe [31–34], Australia [35–37], and others [14,38–41]. These studies conclude that GPM IMERG showed better performances than TRMM products across many countries of the world with different climatic conditions. Most of the SPPs showed weak performance in precipitation detection and resulted in high mean errors in regions having rapid precipitation gradients with complex terrain.

A very few studies are conducted over Pakistan [9–11,42–44]. Cheema and Bastiaanssen [9] calibrated TRMM-TMPA on a monthly time scale over the Indus Basin for runoff and soil water balance studies. Results revealed that TMPA overestimated the precipitation in high ranges and foothill plains, whereas underestimated precipitation in northwest and coastal areas compared with SPG data. Khan et al. [11] conducted a study to assess TMPA-v7 Real Time (RT) and Adjusted (Adj) TMPA and CMORPH-RT during the monsoon season on a daily temporal scale. The result revealed that TMPA-Adj v7 and CMORPH overestimated the precipitation over high altitudes in the study area. Overall, TMPA-v7 performed well as compared to other precipitation products. Anjum et al. [44] assessed the improvements of TMPA-v7 over TMPA-v6 on a basin scale study over the Swat River watershed (14,039 km^2). Results showed that both the magnitude and spatial variation were not captured accurately using TMPA-v7 and v6 on the annual and seasonal scales. However, both the products showed improved accuracy on a monthly time scale than daily precipitation estimate. Hussain et al. [43] evaluated the performance of CMORPH, TMPA, and PERSIANN data sets over three geomorphological climate zones: plain, mountain and glacial regions of Pakistan. Their results indicated that all SPPs captured the precipitation pattern accurately but overestimated the precipitation at glacial, whereas good performance over the mountain region as compared to plain zones. Furthermore, they concluded that the Adjusted (Adj) versions of SPPs outperformed their respective real-time (RT) versions. Muhammad et al. [42] proposed an ensemble algorithm based quantification of precipitation and assigned relative regional performance weights to IMERG research (IR), IMERG real-time (IT), and TRMM 3B42 (RT) to produce regional precipitation (RP) estimates. The results showed that the proposed RP algorithm provided significant agreement with SPG observations. Iqbal and Athar [10] validated TRMM-TMPA with SPG data and Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) named as APH data. The result revealed that TMPA slightly overestimated the precipitation at daily, monthly and annual scales compared with SPG and APH based observations, but the correlation with SPG observation was stronger compared with APH. Based on the seasonal evaluation, TMPA overestimated both SPG and APH based observations during the pre-monsoon and monsoon seasons whereas underestimated during the post-monsoon and winter seasons. Moreover, the performance of TMPA in plain- and medium-elevated regions was better in comparison with high elevated regions.

Generally, in Pakistan, the literature demonstrated that TRMM (RT and Adj) and TMPA performance is better in plain and medium elevated areas, however, both the products overestimate the precipitation in mountainous and glacial regions in the north and coastal areas. Based on the studies conducted over Pakistan, it is concluded that the performance of different SPPs shows significant temporal (daily, monthly, and annual), seasonal (pre-monsoon, monsoon, and post-monsoon), precipitation intensities, climate, and geotopographic dependency. Most of the studies conducted in Pakistan evaluated the TRMM-TMPA, PERSIANN, PERSIANN-CDR, and CMORPH. However, GPM-IMERG is still to be evaluated against other SPPs in Pakistan.

The objectives of this study are twofold: (1) to evaluate the GPM-IMERG with previously evaluated SPPs in Pakistan such as TMPA_3B43 v7, PERSIANN, and PERSIANN-CDR based on climate and topographic diversity and seasonal precipitation variation, i.e., pre- and post-monsoon, monsoon and winter seasons and (2) to select best SPPs and propose an Ensemble Precipitation (EP) algorithm and its evaluation with the selected SPPs in Pakistan. Unlike many other studies available in the literature, a novelty in this study is the development of the EP algorithm by applying the paired sample *t*-test and principal component analysis (PCA) methods.

2. Materials and Methods

2.1. Study Area

Pakistan is situated geographically between 24° and 37°N latitude and 62° and 75°E longitude in the western zone of South Asia and covers approximately 79.6 million ha of the area (Figure 1). Pakistan has complex and diverse topography with elevation ranges from 0 m (Arabian sea) to the world famous Himalayas and Karakoram mountain ranges with a peak elevation of 8600 m [45]. The landscape diversity ranges from glacial and snow cover region in the north, to a plain zone having deserts and plateaus in the middle, and a coastal zone along the Arabian Sea in the south. Due to diverse climatic regimes, the annual precipitation of Pakistan ranges from 300 mm in the south to about 1500 mm in the north. The maximum precipitation occurs during summer (Monsoon) and winter (western disturbances) seasons [46]. Monsoon precipitation occurs during July to September and originates from the Bay of Bengal and enters Pakistan from east and northeast. During the monsoon season, heavy precipitation (from 55 to 60% of annual precipitation) is received in the country [47]. Winter precipitation occurs during December to March due to the Mediterranean Sea and enters Pakistan from Iran and Afghanistan. Moderate precipitation (30% of annual precipitation) is received during the Winter season [48].

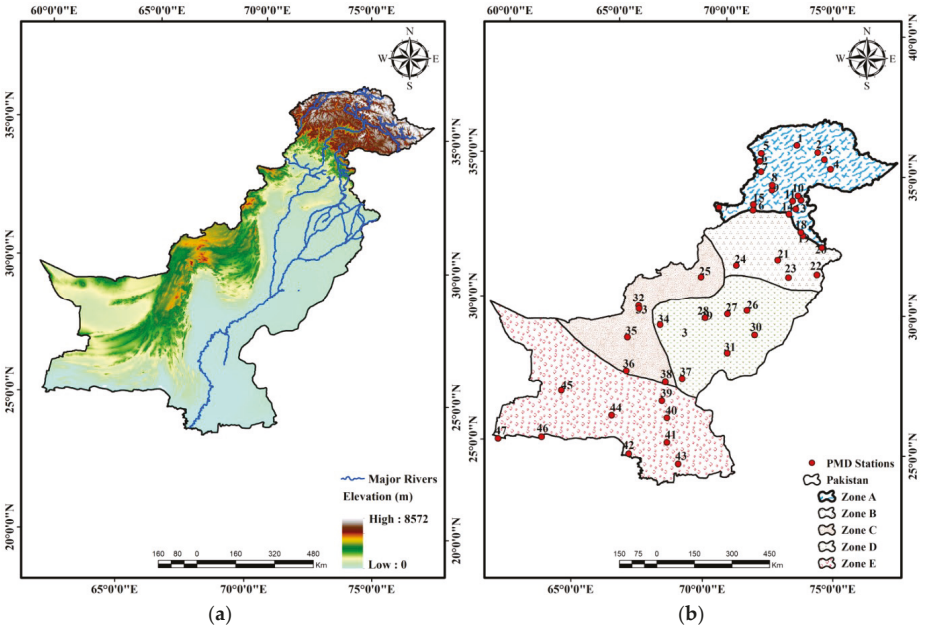


Figure 1. (**a**) Elevation and major rivers of Pakistan. (**b**) Five climate zones of Pakistan and selected Pakistan Meteorological Department (PMD) surface precipitation gauges (SPGs) (the SPG numbers and associated SPG names are shown in Table 1).

The Indus River is the most important river in Pakistan. The Indus River originates from Mount Kailash in Tibet (China) and discharges into the Arabian Sea. The Indus River covers most of Khyber Pakhtunkhwa, Punjab and Sindh provinces of Pakistan, with the Himalayan mountains (North–East) and Highland regions in the north and the arid regions in the southwest. Salma et al. [49] divided Pakistan into five different microclimate zones (Figure 1).

2.2. In Situ Data

Surface Precipitation Gauges (SPGs) are considered as ground truth data, as they provide a direct record of the precipitation at a specific location. The SPP observations at the grid cell are compared with the corresponding SPG during the calibration and validation phases. To monitor the precipitation and climate variation in Pakistan, PMD has established a significant number of weather SPGs all over Pakistan, including some dating back to 1950, some seasonal ones, and some newly established ones. A total of 97 observatories are installed all over Pakistan on behalf of PMD. The precipitation data is collected manually which is subjected to human-induced and instrumental errors. Other associated errors with the SPG located at high elevation regions comes from wind effect, which affects the precipitation measurement by the SPGs. In these circumstances, PMD follows the World Meteorological Organization (WMO) standard code WMO-N for the evaluation and correction of SPG precipitation data to ensure the consistency in the measurements. In this research, daily precipitation data at 47 SPGs are obtained from PMD for the year 2015 and 2016. Table 1 shows the details of all the SPGs utilized in this study.

The dataset is divided into five different climate zones (Zone-A to Zone-E) for evaluation of the topographical and seasonal performance of SPPs. Zone-A is situated from 34°N to 38°N, has cold climate and high mountains like Hindukush, Himalaya, and Koh-e-Sufaid with mean annual precipitation (MAP) of 1034.43 mm. Zone-B lies between 31°N and 34°N and has a mild cold climate

and submountains (MAP of 990.72 mm). Zone-C is located between 27°N and 32°N and has cold weather in winter and hot weather in summer (MAP of 317.00 mm). Zone-D is the hottest and arid zone of Pakistan. The area is almost plain in elevation with some area included in the famous Thar Desert (MAP of 322.41 mm). Zone-E is the coastal zone of Pakistan and situated near the Arabian Sea. The climate in this zone mostly ranges from arid to hyper-arid (MAP of 146.65 mm).

The zero-order method test is used to fill the missing data. The method employs a simple procedure by replacing the average for the missing values [50]. The data quality test is used to check the accuracy and consistency of the data. There are various statistical methods to test the quality of the data. It mainly contains the normality test such as the Jarque–Bera test, histogram (graphical approach), the mean, the skewness, and kurtosis of the data [51]. For the current study, the skewness and kurtosis method is applied to test the data quality.

2.3. Description of Satellite-Based Precipitation Datasets

Three different satellite-based precipitation datasets have been used in the present study. Which include the Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG), Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN).

The IMERG is the quasi-global (between 60°N and 60° S) multi-satellite precipitation product of Global Precipitation Measurement (GPM) [52]. The GPM-IMERG mission was launched in February 2014 and has a spatial resolution of 0.1° × 0.1° and a temporal resolution of 30 min. IMERG is the Level 3 precipitation algorithm of GPM, which has three runs, i.e., early run (latency of 6 h and can be used for warning a probable flood event or landslides), late run (latency of 18 h and suitable for drought monitoring and agricultural forecasting), and final run (latency of three months and used for observation like precipitation observation). In this study, IMERG final run will be used to estimate the seasonal and topographical variation of precipitation in Pakistan.

Tropical Rainfall Measurement Mission (TRMM) is the first space-borne SPP launched in November 1997 its detailed description is available in previous literatures [53–55]. TRMM has three onboard instruments, Precipitation Radar (PR), TRMM Microwave Image (TMI), and Visible Infrared Scanner (VIRS) for recording the data. TMI is a multichannel passive microwave radiometer that supplements the PR by providing total hydrometeor (liquid and ice) contents, with the precipitation systems. The VIRS dispensed the cloud context of precipitation structures and used to connect microwave precipitation information to infrared-based precipitation estimates from geosynchronous satellites [53–55]. Currently, the Multi-satellite Precipitation Analysis (TMPA) is producing the best precipitation estimates at 0.25° spatial resolution for the areas between 50°S and 50°N. The TMPA algorithm combines the precipitation estimates from various satellite systems, as well as SPG precipitation analyses [17,56]. TRMM TMPA-3B43 v7, a daily precipitation product, is used in this study.

The PERSIANN is an algorithm developed by the Center for Hydrometeorology and Remote Sensing (CHRS) in the University of California Irvine with a spatial resolution of 0.25° × 0.25°. In PERSIANN PMW data (TMI, AMSU-B, and SSM/I) is used to adjust the neural network parameters to increase precipitation estimation accuracy. PERSIAN–CDR is another product from PERSIANN family developed by CHRS with the spatial resolution of 0.25° × 0.25°. PERSIANN–CDR uses the same neural network for precipitation estimation as used previously but differs regarding input IR dataset with the use of GridSat-B1 instead of CPC-IR. Moreover, PMWs data are not used [57]. In this study, PERSIANN and PERSIANN-CDR algorithms are evaluated based on seasonal and regional scales.

Table 1. PMD SPGs used in the study along with SPG number (shown in Figure 1b), Latitude (N), Longitude (E), and Elevation (m).

SPG Number	SPG	Latitude (°)	Longitude (°)	Elevation (m)	SPG Number	SPG	Latitude (°)	Longitude (°)	Elevation (m)
1	Gupis	36.17	73.40	2156.00	26	Multan	30.20	71.43	122.00
2	Gilgit	35.92	74.33	1468.00	27	D.G.Khan	30.05	70.63	148.00
3	Bunji	35.67	74.63	1372.00	28	Chhor	29.80	69.70	5.00
4	Astore	35.37	74.90	490.43	29	Barkhan	29.88	69.72	1097.00
5	Chitral	35.85	71.83	1498.00	30	Bahawal Pur	29.39	71.68	110.00
6	Drosh	35.57	71.78	1464.00	31	Khanpur	28.80	70.40	89.00
7	Dir	35.20	71.85	1375.00	32	Quetta	30.18	66.95	1719.00
8	Saidu Sharif	34.73	72.35	961.00	33	Quetta (Samungli)	30.08	66.97	1626.00
9	Balakot	34.55	72.35	1528.35	34	Sibi	29.55	67.88	133.00
10	Muzaffarabad	34.37	73.48	838.00	35	Kalat	29.02	66.58	2015.00
11	Kakul	34.18	73.25	1308.00	36	Khuzdar	27.83	66.63	1231.00
12	Murree	33.92	73.38	2291.00	37	Rohri	27.67	68.90	66.00
13	Islamabad	33.58	73.09	507.00	38	Larkana	27.53	68.23	53.00
14	Garhi dupatta	34.22	73.62	814.00	39	Padidan	26.85	68.13	46.00
15	Peshawar	34.50	71.30	327.00	40	Nawab Shah	26.25	68.37	37.00
16	Cherat	33.82	71.89	1372.00	41	Hyderabad	25.38	68.41	28.00
17	Parachinar	33.86	70.08	1775.00	42	Karachi Airport	24.91	66.93	22.00
18	Mangla	33.07	73.63	284.00	43	Badin	24.63	68.90	9.00
19	Jhelum	32.90	73.30	288.00	44	Lasbella	26.23	66.17	87.00
20	Sialkot	32.50	74.50	255.00	45	Panjgur	26.97	64.10	968.00
21	Sarghoda	32.10	72.60	187.00	46	Pasni	25.27	63.48	9.00
22	Lahore PBO	31.60	74.40	216.00	47	Jiwani	25.07	61.80	56.00
23	Faisal Abad	31.34	73.13	185.00					
24	D.I.Khan	31.80	70.60	173.00					
25	Zhob	31.35	69.47	1405.00					

2.4. Adjustments in Spatial and Temporal Resolutions

All of the SPPs used in this study differ in spatial and temporal resolutions. To adjust the spatial resolution, an aggregation method was used to aggregate the precipitation product of IMERG from 0.1° to 0.25°, and the areal aggregation weights of 0.16, 0.08, and 0.04 were assigned to four IMERG grid cells falling inside a 0.25° TRMM grid cell, four located halfway within, and the ninth covers one-fourth inside, respectively [40]. Furthermore, the coordinate matching process was also performed to avoid any mismatch issue. In the coordinate matching process, the precipitation record and the latitude/longitude of the SPGs are matched against the amount of satellite precipitation in the pixel. To produce a single mean value at each pixel, the precipitation data is averaged when a pixel has more than one SPG.

The day in Pakistan begins 5 h ahead of Greenwich Mean Time (GMT+5). The satellite-based data accumulations are computed from 3:00 to 3:00 UTC to match the 8:00 to 8:00 local time of the SPG data in Pakistan. In order to compare the SPPs estimates with the SPGs data, the satellite precipitation data is accumulated into daily precipitation, which has the same temporal resolution as the SPG precipitation data.

2.5. Precipitation Products Evaluation

The selected four SPPs are quantitatively evaluated by comparing the SPPs with the daily PMD SPG observations from January 2015 to December 2016. The statistical indices used to assess the efficacy of selected SPPs are listed in Table 2, including (i) Mean Error (ME) to evaluate the bias, (ii) Root Mean Square Error (RMSE) to evaluate the magnitude of error with more emphasis on large errors as compared to ME, and (iii) the Pearson correlation coefficient (CC). For comprehensive evaluation of CC, the CC ranges mentioned by Iqbal and Athar [10] are considered: Weak (CC < 0.25), Low (0.25 < CC < 0.5), Moderate (0.5 < CC < 0.75), and Strong (0.75 < CC < 1).

Table 2. List of metrics used in statistical performance measures [a].

Statistical Measures	Equation	Perfect Value
Mean Error (ME)	$\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)$	0
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2}$	0
Correlation Coefficient (CC)	$\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$	1
Probability of Detection (POD)	$\frac{hit}{hit+miss}$	1
False Alarm Ratio (FAR)	$\frac{false\_alarm}{hit+false\_alarm}$	0
Critical Success Index (CSI)	$\frac{hit}{hit+miss+false\_alarm}$	1

[a] X_i and Y_i represents SPP- and SPG-based precipitation observations for the ith time step, respectively, $\overline{X}$ and $\overline{Y}$ are the average values of SPP- and SPG-based observations, n represents the sample size, *hit* is the number of events when the precipitation is recorded both by SPPs and SPGs, *miss* is the number of events missed by SPP and recorded by SPG observations, *false_alarm* is the number of events that SPP capture the precipitation while SPGs show no precipitation record.

For more accurate quantification of errors, three additional categorical statistical indices, including Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI), are used to detect the agreement between SPP- and SPG-based precipitation data. POD indicates how well the SPPs detect the precipitation events among all the SPG-based precipitation events (1 mm was used as the precipitation/no precipitation threshold [22], FAR indicates the fraction of incorrectly detected

precipitation events by SPPs, and CSI represents an overall fraction of precipitation events correctly detected by the SPPs.

2.6. Algorithm for Ensemble Precipitation Estimation

The proposed algorithm for satellite-based ensemble precipitation estimation is based on K-fold cross-validation, paired sample *t*-test, Principal Component Analysis (PCA), and the ensemble algorithm.

The K-fold cross-validation method is used to evaluate the SPPs performance and its selection. In the K-fold cross-validation, the SPG observations in the five different climate zones are divided into K-datasets. Furthermore, these datasets are partitioned into training and testing sets. During the validation process, the data of a single SPG (testing set) in a selected climate zone is taken out from the calibration dataset (training set). The average error across all K-trials was calculated. The K-fold test was repeated for all the SPGs considered in the study.

A dependent-sample *t*-test, also known as paired sample *t*-test, is used to compare the SPPs with the SPG datasets. The paired sample *t*-test assesses whether the mean difference between paired objects/observations is significantly different from zero. The following two hypotheses are tested for the paired sample *t*-test.

Hypothesis 1. *Null Hypothesis*:

$$H_0 : \mu_1 = \mu_2 \tag{1}$$

Hypothesis 2. *Alternative (Nondirectional) Hypothesis*:

$$H_a : \mu_1 \neq \mu_2 \tag{2}$$

Hypothesis 3. *Alternative (Directional) Hypothesis*:

$$H_a : \mu_1 < \mu_2 \text{ or } H_a : \mu_1 > \mu_2 \tag{3}$$

where μ_1 and μ_2 are the mean of first and second variables. Moreover, the Wilcoxon matched-paired signed ranks test is also performed to examine the extreme violation of normality assumptions and further validation of the performance of SPPs. The mean error and standard deviation between the SPP and SPG datasets are calculated. Based on K-fold cross-validation and paired sample *t*-test with 95% confidence interval, the best satellite products (minimum mean error and standard deviation) are selected.

The Principal Component Analysis (PCA) technique uses a sophisticated mathematical principle to transform a number of possibly correlated variables into a smaller number of variables called principal components. The basic idea behind the application of PCA is to select the two best SPPs out of the four SPPs and to determine the percentage correlation of the selected SPPs with the ground truth data (PMD SPGs). This percentage is considered as a weight factor "w" ($\Sigma w = 1$) of the particular satellite product.

After calculating the weights, the precipitation observations of the selected SPPs over each climate zone have been combined using the following equation:

$$EP_i = w_1 S_i + w_2 s_i,\ i = 1, 2, \ldots, N \tag{4}$$

where EP is the ensemble precipitation, w_1 and w_2 are weights associated of the selected two SPPs, S and s are the two selected satellite products, i is the number of PMD SPGs, and N is the maximum number of SPP records in the particular climate zone.

Finally, the developed ensemble precipitation product EP = [EP_1, EP_2, . . . , EP_N] is compared and evaluated with the corresponding ground observations G = [G_1, G_2, . . . , G_N] using the statistical metrics presented in Table 2.

3. Results

The results obtained are divided into the following three sections. Section 3.1 describes the regional evaluation of IMERG, TMPA, PERSIANN, and PERSIANN-CDR, Section 3.2 describes the evaluation of ensemble precipitation (EP) algorithm, and Section 3.3 describes the validation of satellite products and EP algorithm on seasonal scales.

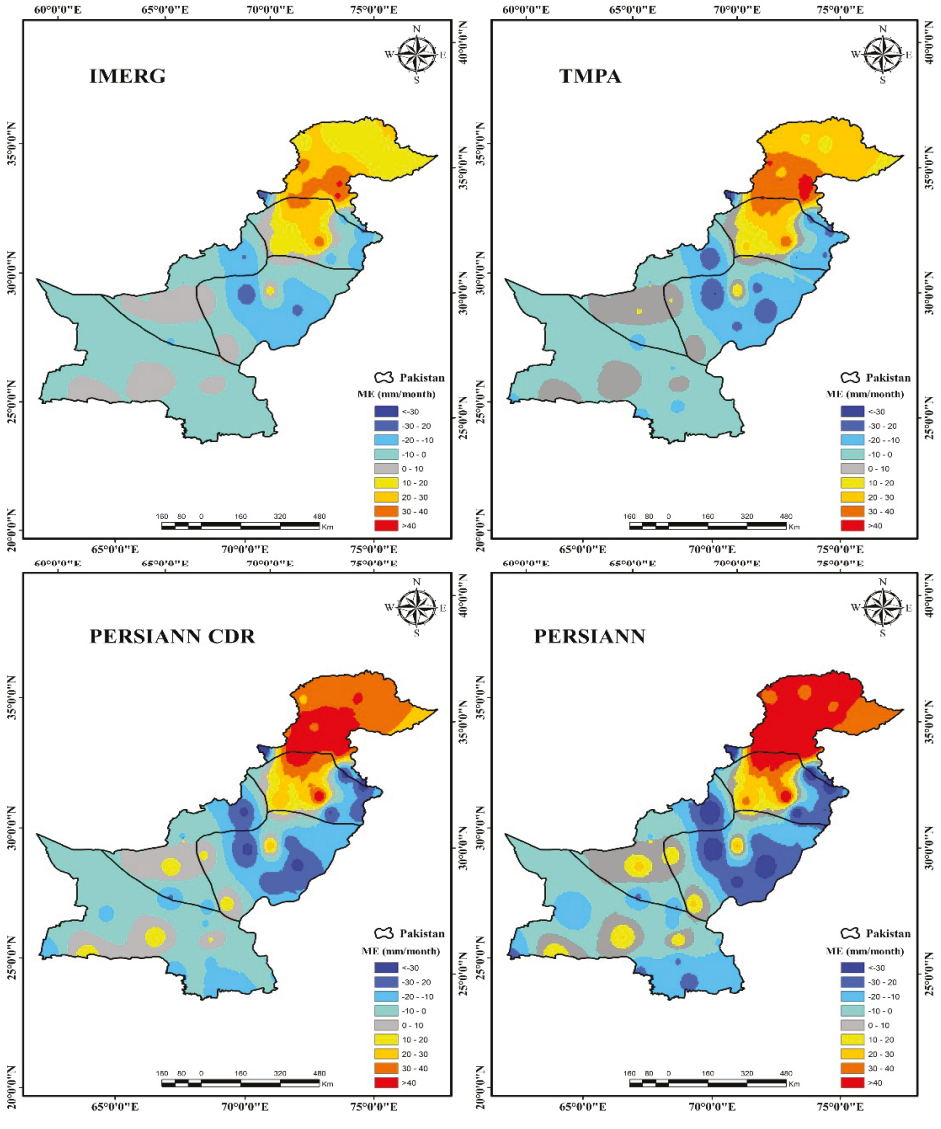
3.1. Regional Evaluation of Satellite Products

Accurate estimation of precipitation distribution is a challenging task over the mountainous region because of high precipitation variability due to the orographic enhancement and terrain-induced errors on remote sensing retrievals. The northern regions of Pakistan also pose uncertainties in precipitation estimation due to complex climatology and terrain of the region. Therefore, this research is conducted to test the abilities of IMERG over the complex topography of Pakistan by comparing with other SPPs (TMPA, PERSIANN, and PERSIANN-CDR) and against the SPG data. The regional performance of all the SPPs is evaluated on the annual temporal scale at different climate regions using statistical metrics listed in Table 2. The regional variability of all these metrics for IMERG, TMPA, PERSIANN, and PERSIANN-CDR over the selected five climatic regions is shown in Figure 2.

The ME results (Figure 2a) indicated that all the SPPs performed poorly in Zone-A (considerable overestimation of precipitation except at Parachinar SPG) as compared to the other zones. This error may be associated with the fact that in the glacial region, the presence of ice cover and cold cloud contribute to the overestimation [58,59]. Generally, the magnitude of ME shows a decreasing trend from North to South of the country. All the satellite performed well in Zone-E (slight underestimation) which is coastline area with relatively low elevation and rainfall intensity. Moreover, the performance of all SPPs is good in the Zone-C and Zone-B while moderate in Zone-D. Among all the SPPs, IMERG product showed significantly improved performance followed by TMPA. Higher ME (ME > 40 mm/month) is observed over Murree, Muzaffarabad, Dir, Drosh, etc., while smallest ME is observed at Jiwani and Badin SPGs.

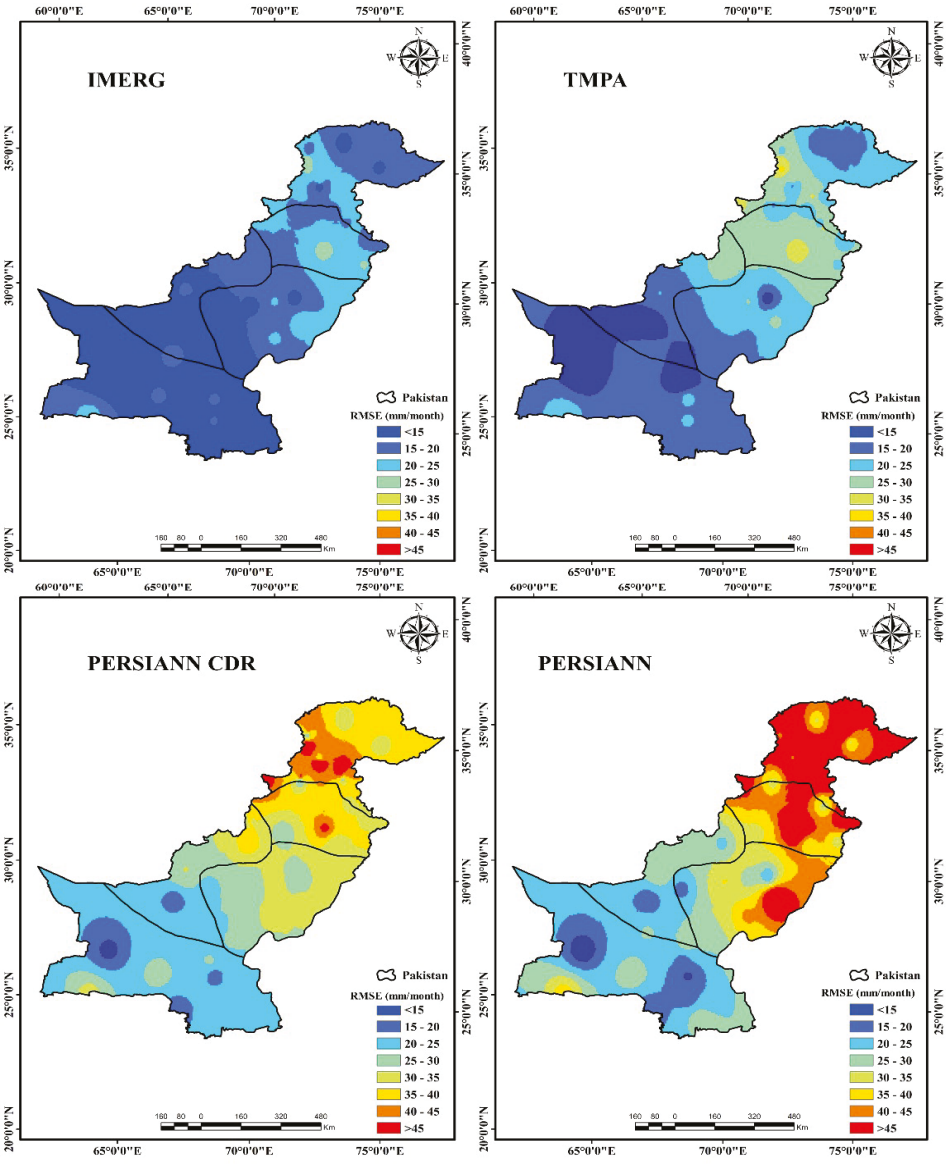
Figure 2b indicates that RMSE is showing a similar trend to ME: higher RMSE at Zone-A and lower RMSE at low altitude climate zones. However, the RMSE of Zone-D is slightly higher compared to Zone-C and Zone-D. Maximum RMSE is observed at Drosh, Dir, and Balakot while minimum RMSE is observed at Larkana and Panjgur SPGs. Considering the Zone-D (a special scenario), maximum RMSE is observed at SPGs namely Khanpur and Bahawalpur, revealing the poor performance of satellite SPPs at those SPGs. Similar observations are found in the study conducted by Yawar [43], where RMSE is decreasing from North to South. The maximum RMSE associated to IMERG is 29.23 mm/month at Drosh SPG, while for TMPA the maximum RMSE is 38.50 mm/month at Dir SPG. For PERSIANN and PERSIANN-CDR maximum RMSE is observed Dir (65.14 mm/month) and Balakot (55.9 mm/month), respectively.

The performance evaluation based on CC (Figure 2c) revealed that IMERG agreed well with the majority of SPG observations as compared to TMPA, PERSIANN, and PERSIAN-CDR. The results revealed that IMERG performance is good over the Zones-B, C, and D, and average performance over Zone-A and Zone-E. IMERG shows high correlation with Bahawalpur SPG in Zone-D followed by Karachi airport (Zone-E), and Faisalabad (Zone-B) while relatively poor performance at Lasbella (Zone-E) and Sialkot (Zone-B) SPGs. However, TMPA has a high correlation with Khanpur (Zone-D), Faisalabad (Zone-B), and Karachi airport (Zone-E) SPGs while lower performance in Lasbella (Zone-E), Sialkot (Zone-B), and Dir (Zone-A). Overall, the results indicated that the performance of IMERG outperformed all other satellites over most of the SPGs.



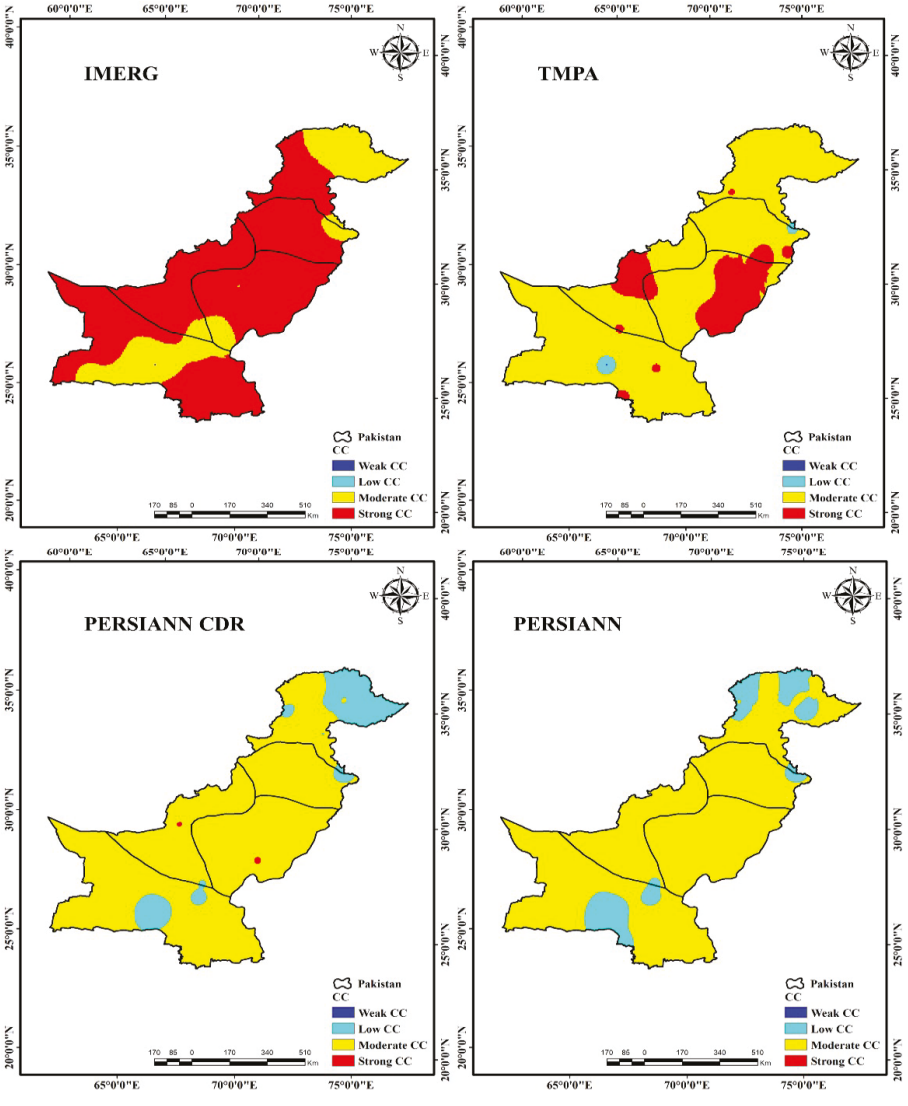
(**a**)

Figure 2. *Cont.*



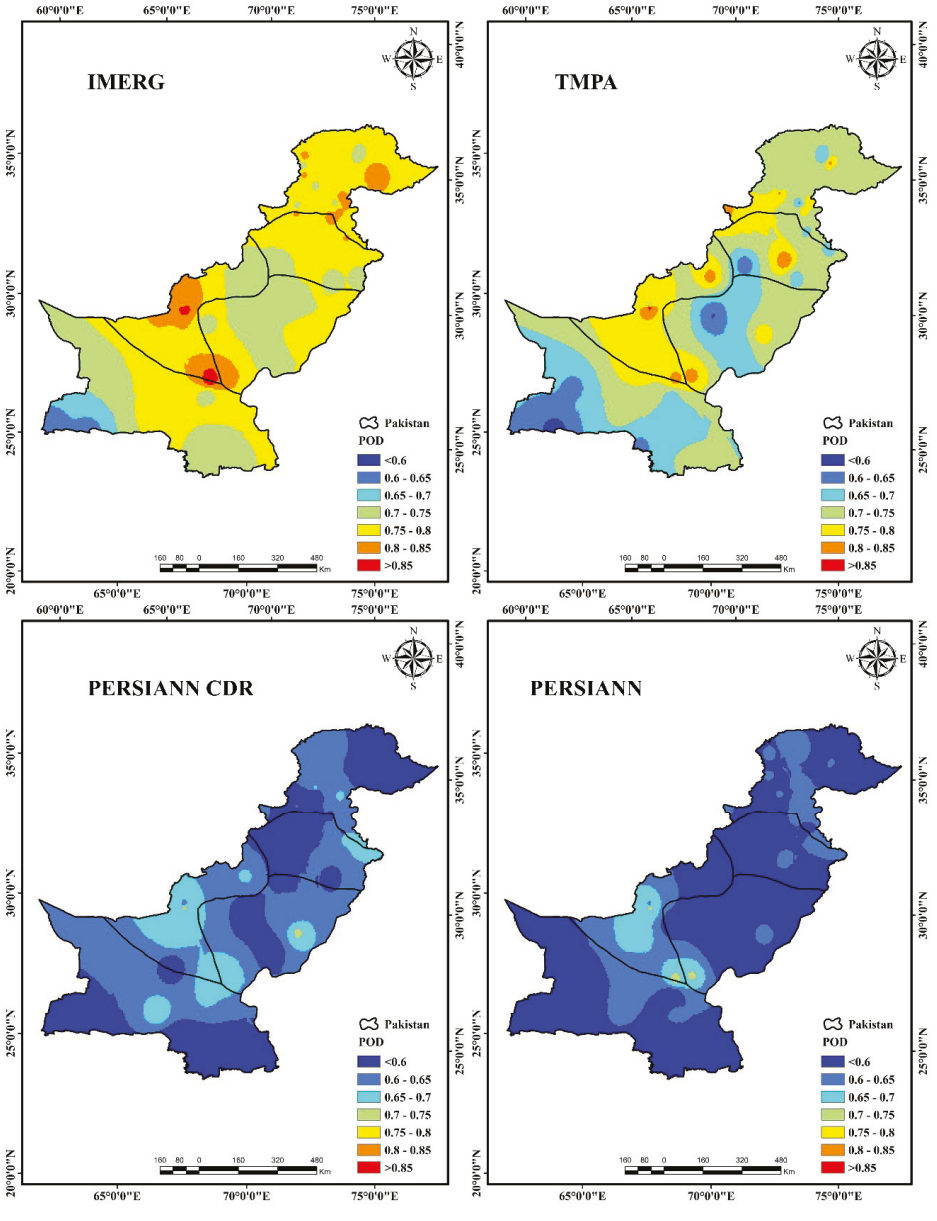
(**b**)

Figure 2. *Cont.*



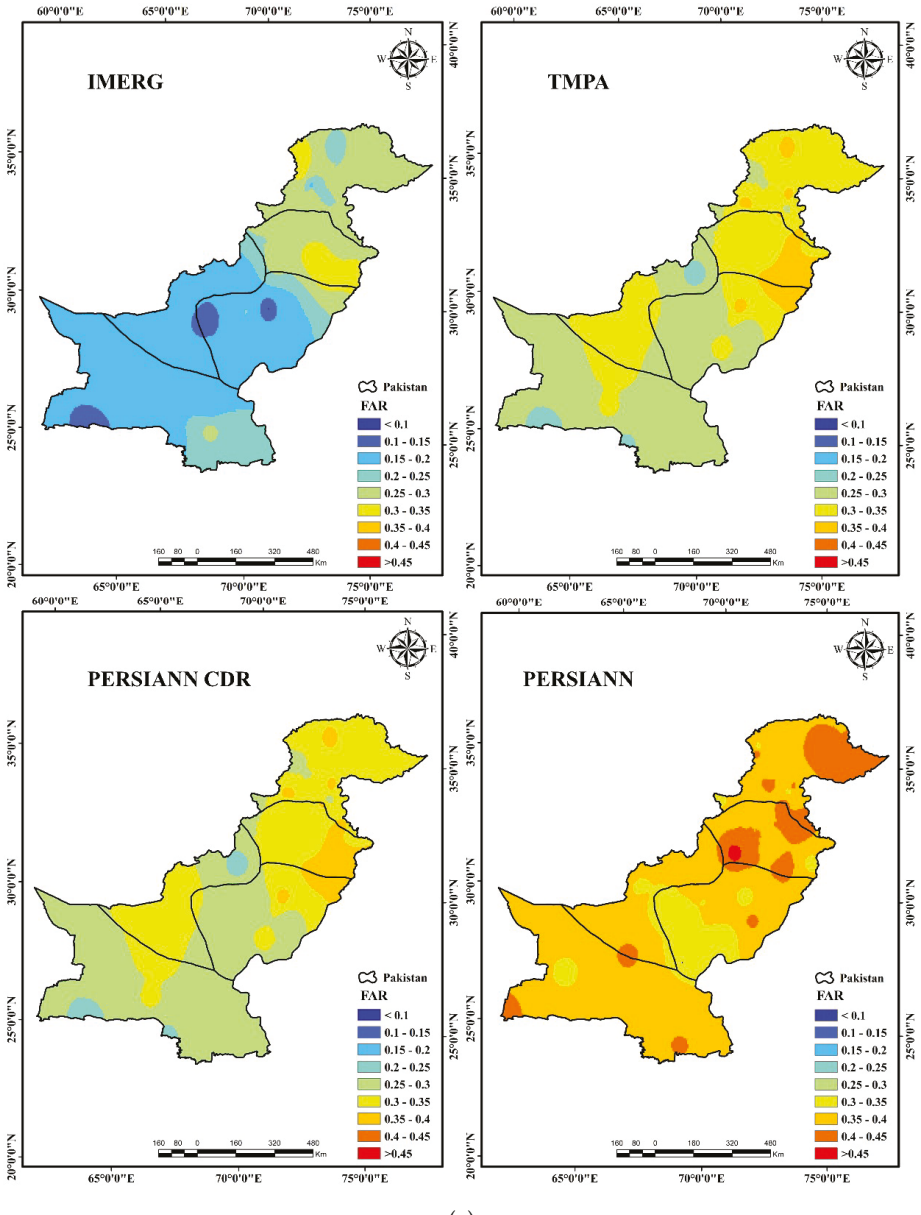
(**c**)

Figure 2. *Cont.*



(**d**)

Figure 2. *Cont.*



(**e**)

Figure 2. (**a**) Regional distribution of mean error (ME) of satellites precipitation estimates. (**b**) Regional distribution of RMSE of satellite precipitation estimates. (**c**) Regional distribution of CC of satellite precipitation estimates. (**d**) Regional distribution of POD of satellite precipitation estimates. (**e**) Regional distribution of FAR of satellite precipitation estimates.

Figure 2d,e shows the POD and FAR (evaluated on daily time scale) of all the satellite SPPs, respectively. The POD of IMERG and TMPA are significantly comparable over most of the SPGs in

all the climate zones; whereas, the POD of PERSIANN and PERSIANN-CDR is slightly lower over most of the SPGs. However, FAR distribution of all the precipitation products is completely different from each other. In the case of IMERG, lower FAR values are observed over the plain regions as compared to TMPA. The CSI of all the precipitation products exhibits almost the similar trend, and the zone-based details are presented in Table 3.

Table 3. The average CSI of SPPs over the five climate zones.

SPPs	Zone-A	Zone-B	Zone-C	Zone-D	Zone-E
		Critical Success Index (CSI)			
IMERG	0.28	0.34	0.30	0.37	0.30
TMPA	0.23	0.27	0.24	0.28	0.26
PERSIANN	0.17	0.23	0.21	0.24	0.22
PERSIANN-CDR	0.21	0.23	0.22	0.26	0.22

Overall, the results presented that all the SPPs show the elevation dependencies which are characterized by overestimation at higher and slight/moderate underestimation over lower altitude regions. The IMERG performed best followed by TMPA at the majority SPGs in all zones. Furthermore, the SPPs are evaluated with regional SPGs, and high correlations of IMERG and TMPA with SPG observations are found. The performance of PERSIANN and PERSIANN CDR is found relatively poor as compared to IMERG and TMPA.

3.2. Ensemble Precipitation (EP) Algorithm Evaluation

Tables 4 and 5 show the paired sample *t*-test and PCA results, respectively. The paired sample *t*-test result shows the lowest mean error and standard deviation between the PMD and IMERG, followed by PMD and TMPA. This lowest mean error and the standard deviation are taken under consideration for EP algorithm which represent higher correlation between the compared pairs. The standard deviation and mean error of PMD-IMERG and PMD-TMPA is considerably low as compared to PMD-PERSIANN and PMD-PERSIANN CDR. Based on the high correlation, low standard deviation and mean error, IMERG, and TMPA satellite products are selected for the EP algorithm. The results of PCA, shown in Table 5, shows that the high correlation was found between the PMD-IMERG (61.82%) and PMD-TMPA (23.83%). Since the PERSIANN and PERSIANN CDR satellite products' performance was not satisfactory, their percentages and associated weights are distributed among the IMERG and TMPA for the final EP algorithm. The final weights for IMERG and TMPA are calculated by linear correlation, and the magnitude is 0.71 and 0.29, respectively.

Table 4. Paired sample *t*-test of satellite products over the five climate zones.

Pairs	Mean	Standard Deviation	Standard Mean Error	95% Confidence Interval of the Difference	
				Lower	Upper
PMD-IMERG	1.26	13.83	2.02	−2.80	5.32
PMD-TMPA	2.99	18.92	2.76	−2.55	8.55
PMD-PERSIANN	10.93	29.92	4.36	2.14	19.71
PMD–PERSIANN CDR	2.31	27.05	3.94	−5.63	10.25

Table 5. Principal component analysis (PCA) of satellite products over the five climate zones.

Satellite Products	Eigenvalues	Percentage Contribution	Percentage Weight
IMERG	2.47	61.82	0.62
TMPA	0.86	23.83	0.24
PERSIANN	0.20	5.18	0.05
PERSIANN CDR	0.44	9.17	0.09

The developed EP algorithm framework was evaluated using the same statistical metrics presented in Table 2. Table 6 shows the complete details of the Statistical Metrics for the EP algorithm at all the PMD SPGs. The result (Table 6 and Figure 3) indicated that ME and RMSE are significantly decreased while CC increased at most of the SPGs as compared to IMERG, TMPA, PERSIANN, and PERSIANN-CDR. However, some of the SPGs such as Cherat, Garhi Dupatta, Kakul, and Muzaffarabad still have higher ME and RMSE.

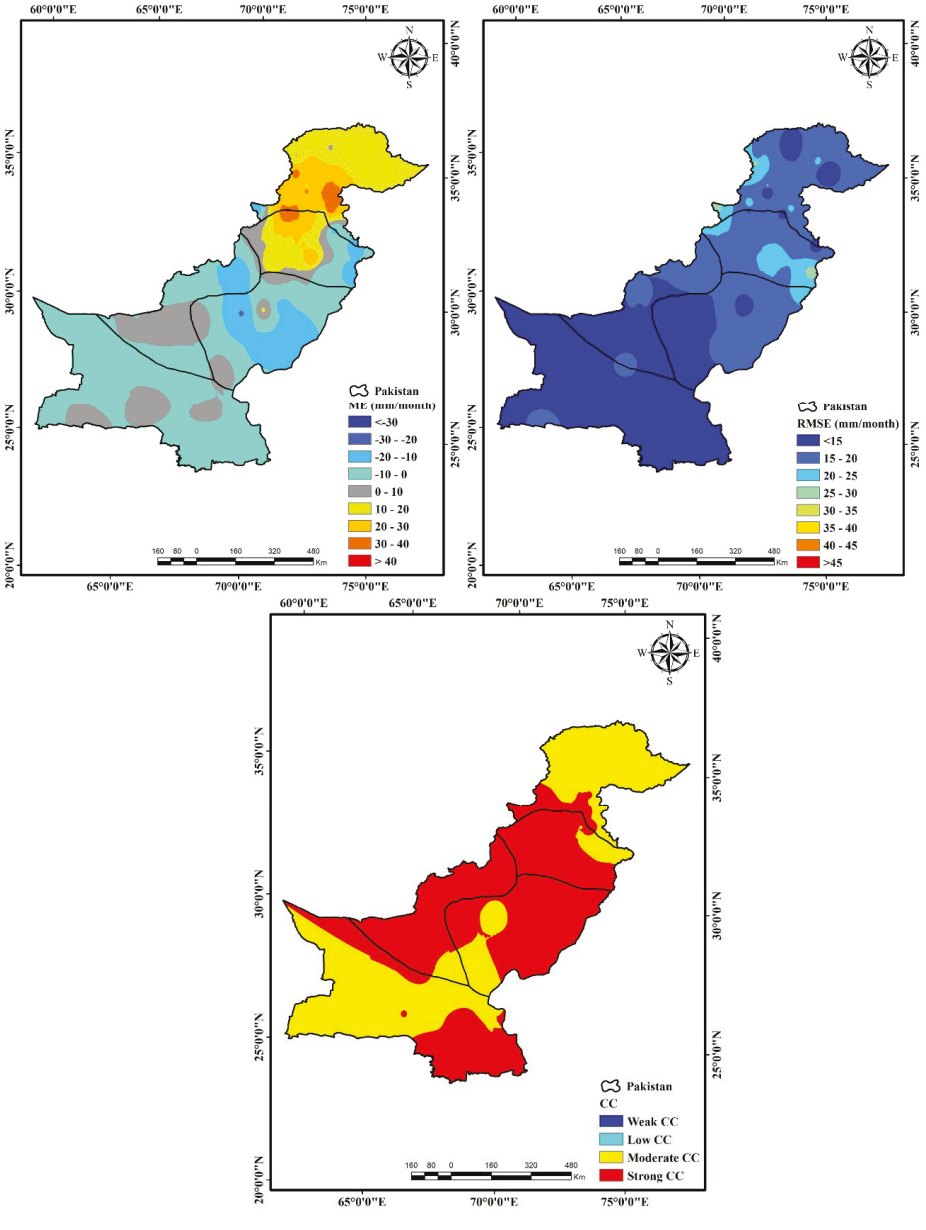


Figure 3. EP algorithm evaluation.

Table 6. Statistical evaluation of ensemble precipitation (EP) algorithm at selected SPGs.

SPG	ME (mm/month)	RMSE (mm/month)	CC	SPG	ME (mm/month)	RMSE (mm/month)	CC
Badin	−4.29	9.06	0.91	Kakul	29.89	19.58	0.82
Chhor	−21.71	12.86	0.82	Kalat	5.73	8.19	0.92
Hyderabad	−1.15	11.86	0.82	Khanpur	−14.61	18.49	0.87
Karachi Airport	−5.69	9.75	0.91	Khuzdar	−9.84	19.31	0.91
Astore	18.34	9.60	0.78	Lasbella	5.01	7.11	0.62
Balakot	31.66	9.28	0.84	Muzaffarabad	39.95	20.00	0.87
Cherat	39.86	13.80	0.82	Nawab Shah	4.58	8.33	0.87
Chitral	9.44	18.07	0.84	Padidan	−5.29	9.35	0.65
Dir	33.79	23.12	0.84	Panjgur	−3.68	4.99	0.82
Drosh	19.95	26.49	0.82	Parachinar	−19.18	26.11	0.77
Faisal Abad	−9.74	19.42	0.91	Pasni	3.51	18.45	0.77
Garhi Dupatta	32.93	19.27	0.82	Peshawar	24.46	22.27	0.87
Gilgit	12.40	18.20	0.72	Quetta	−9.12	18.83	0.92
Gopis	9.46	9.40	0.83	Rohri	6.49	9.38	0.72
Jhelum	−8.70	19.82	0.72	Sarghoda	29.12	23.51	0.82
Lahore PBO	−15.56	26.23	0.91	Sibi	7.47	13.67	0.82
Multan	−12.90	12.17	0.91	Zhob	−19.34	19.18	0.77
Murree	39.73	21.40	0.81	D.G.Khan	11.01	19.52	0.82
Saidu Sharif	18.51	20.76	0.82	Larkana	−2.25	3.37	0.72
Sialkot	−18.70	13.03	0.62	Mangla	−13.79	16.75	0.82
Bahawal Pur	−18.94	19.74	0.91	Quetta	8.73	13.10	0.92
Barkhan	−19.09	9.64	0.77	(Samungli)			
Bunji	15.44	21.51	0.72	**Maximum**	39.95	26.49	0.92
D.I.Khan	16.84	18.29	0.91	**Median**	5.01	18.07	0.82
Islamabad	13.50	12.02	0.82	**Mean**	5.32	15.69	0.82
Jiwani	−5.59	13.35	0.72	**Minimum**	−21.71	3.37	0.62

3.3. Seasonal Evaluation of Satellite Products and EP Algorithm

The selected SPPs, i.e., IMERG, TMPA, and EP algorithm, were evaluated for the four seasons: pre-monsoon (April, May, and June), monsoon (July, August, and September), post-monsoon (October and November), and winter (December, January, February, and March) over Pakistan. Pakistan receives a higher percentage (60%) of precipitation during the Monsoon season (July to September). The monsoon precipitation varies spatially in magnitude from low (around 100 mm) in the south (24–28°N), higher (>700 mm) in the northeast (29–33°N), and again low (<100 mm) in the far north (glacial region, 34–36°N) [11].

Figure 4 shows the average precipitation where PMD and SPG records are compared with the selected SPPs and developed EP algorithm over the five climate zones in the four seasons. The figure shows that high precipitation is received in the monsoon season followed by the pre-monsoon season. Zone A receive high precipitation, and a decreasing trend is observed as we go toward the south. During the pre-monsoon season, IMERGE significantly overestimated the precipitation in Zone-E. While it significantly underestimated it in Zone-C. Contrarily, TMPA overestimated the precipitation over all the climate zones except Zone-C. However, the EP algorithm showed a better performance in plain areas while promising performance in elevated zones.

Moreover, during the Monsoon season, IMERG closely predicted the precipitation in all the climate zones while TMPA highly overestimated the heavy precipitation in Zone-A and other plain areas and underestimated the precipitation in Zone-C. The EP algorithm slightly overestimated the precipitation in Zone-A while performed well in plain areas. In the post-monsoon season, both the IMERG and TMPA underestimated the precipitation across all the five climate zones except Zone-D and Zone-E. Moreover, EP algorithm closely predicted the precipitation in the plain area, however, overestimated the precipitation at elevated zones. While in the winter season, IMERG and TMPA underestimated the precipitation in all climate zones except Zone-E and Zone-C, respectively. On the other hand, the EP algorithm performed poorly in Zone-C while high performance is observed over other four climate regions. Overall, IMERG showed high performance in capturing heavy precipitation and relative lower performance in moderate/low precipitation seasons. The EP algorithm performed very well in plain areas irrespective

of the season and showed relatively poor performance over mountainous regions. TMPA highly overestimated the heavy precipitation and underestimated the moderate/lower precipitation.

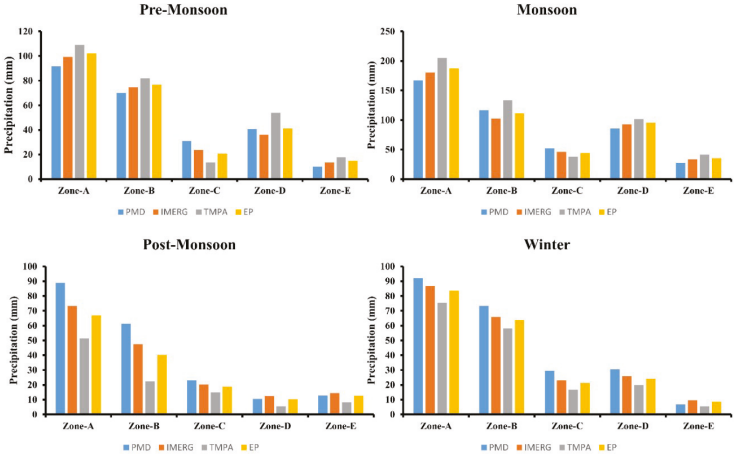


Figure 4. Comparison of mean seasonal (daily accumulated) precipitation over the five climate zones.

The mean error (ME) and root mean square error (RMSE) over the five zones compared to the PMD SPG data is shown in Figures 5 and 6, respectively. Figure 5 depicts that the IMERG, TMPA and EP algorithm overestimated the precipitation in pre-monsoon and monsoon seasons with the exception in Zone-D where IMERG and TMPA underestimate the precipitation. While they underestimated the precipitation in post-monsoon and winter. Based on the magnitude of the ME, the TMPA experienced a high error while the EP algorithm performed well in all the climate zones in every season. Figure 6 presents the magnitude of seasonal variation of RMSE of all the four seasons over the five climate zones. Higher RMSE is observed during the monsoon season followed by pre-monsoon. Generally, the performance of the EP algorithm and IMERG are reasonably well all over the seasons as compared to TMPA.

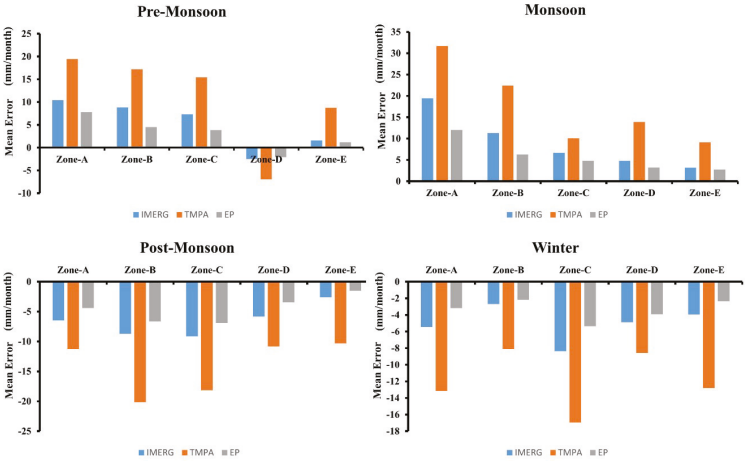


Figure 5. Mean error (ME) of satellite products and EP algorithm during the pre-monsoon, monsoon, post-monsoon and winter seasons.

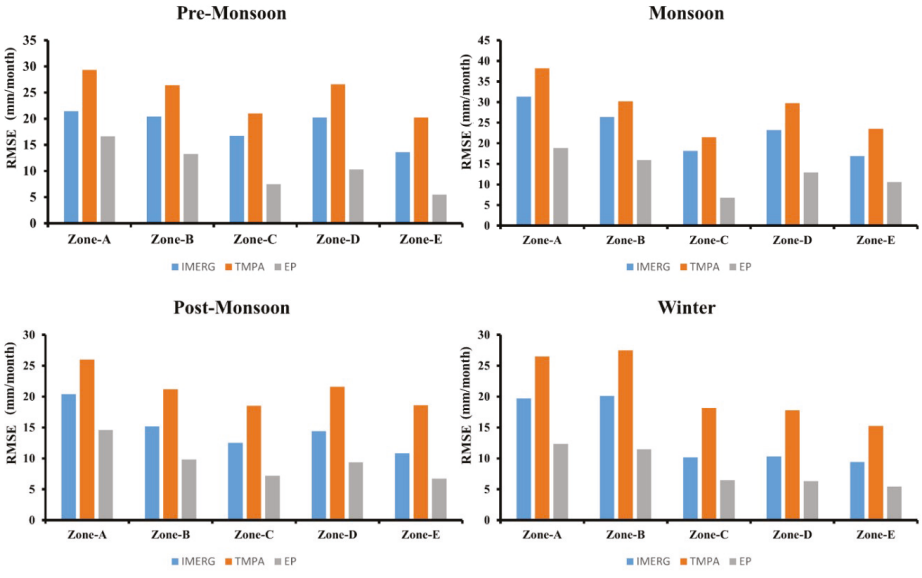


Figure 6. Root mean square error (RMSE) of satellite products and EP algorithm during the pre-monsoon, monsoon, post-monsoon, and winter seasons.

4. Discussion and Conclusions

Accurate estimation of precipitation with high spatial and temporal resolution is a very important for different hydrological simulations and climate change studies. Recently SPPs are attracting the attention of researchers and play a vital role in estimating precipitation in ungauged or poorly gauged regions. However, it is a challenging task in the developing countries such as Pakistan having sparse SPG network and complex topography. Based on the spatial (regional) scale evaluation (Figure 2a–e), our study supports the findings of previous literatures [10,43] conducted in Pakistan. In this study, it has been observed that high errors are found over high elevations, and these errors are reducing with the elevations from north to south. The IMERG and TMPA resulted in a poor performance at high elevated zone (Zone-A) of the study area. There could be number of reasons such as external error associated to the SPG (for example, wind effect, splashing effect of precipitation, evaporation from the SPGs and human-induced errors), complex topography, climate variability in the region, seasonality (pre-monsoon, monsoon, post-monsoon, and winter etc.) [10,13,39], and a sparse SPG network. However, a strong correlation has been observed over the plain elevation and low positive correlation over high elevation. These findings are consistent with the previous studies. However, the EP performed exceptionally well in plain areas while its performance is reducing with the elevation.

The IMERG and TMPA overestimated the precipitation during the pre-monsoon and monsoon seasons while underestimated the precipitation during post-monsoon and winter seasons. Moreover, good agreement between IMERG and SPGs was found in plain and medium elevated regions. These findings are consistent with the previous studies. However, EP performed very well during high and moderate precipitation while its performed ordinary in low precipitation events.

Different algorithms such as algorithms for improved calibration, reducing sampling issue, moving from TRMM to GPM, etc., have been implemented to reduce the non-negligible error in the SPPs. However, there is still room for further advancement in those algorithms to provide consistent results [14,40]. Effort has been made by Muhammad Waseem et al. [42] to minimize the inconsistency issues and the associated errors by merging different SPPs. Their algorithm is based on leave-one-out cross-validation (LOOCV), regional performance weights (MSE_r^j), and the ensemble

algorithm. They concluded that their developed algorithm presented better agreement than the selected SPPs. In this study, we developed ensemble precipitation (EP) product for each SPG in the study area based on selecting two (best) out of four SPPs, i.e., IMERG and TMPA 3B43 v7. The best SPPs were selected using the paired sample *t*-test, and the corresponding weights are calculated using the principal component analysis (PCA). The developed EP algorithm was evaluated based on regional and temporal scales. The performance of the developed EP algorithm based on statistical evaluation is found comparatively better than the RP algorithm developed by Muhammad Waseem et al. [42]. Moreover, the EP algorithm has also been tested on a seasonal scale, and the algorithm performed well when evaluated by ME and RMSE. Overall, the developed EP algorithm performed very well across all the Pakistan irrespective of the seasons.

The EP algorithm is evaluated on spatial and temporal scales, and it has been observed that it outperformed the IMERG and TMPA in the plain and medium elevated areas. However, it performed relatively poorer over high elevated regions. The EP algorithm performed significantly better when evaluated based on temporal (seasonal) scale. Overall, the EP algorithm can capture the spatial precipitation pattern over the region at annual and seasonal scales very well. Moreover, the methodology presented in this study is very simple, and it has the capability to select the best SPPs for a specific region to minimize the errors associated SPPs to ensure consistency in the performance.

Author Contributions: U.R.K. and S.S. formulated the method; U.R.K. and S.M. analyzed the data; U.R.K. and S.S. interpreted the results; U.R.K. wrote the paper; and L.J. and S.M. provided the reviews and suggestions.

Funding: This research was partially supported by the National Natural Science Foundation of China (grant number 51479090).

Acknowledgments: The authors would like to acknowledge the Pakistan Meteorological Department (PMD) for providing the ground truth data and grateful to SPP datasets developers. The valuable comments from the editor and the two anonymous reviewers are greatly appreciated.

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

References

1. Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]
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. 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]
4. 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]
5. Wu, H.; Adler, R.F.; Hong, Y.; Tian, Y.; Policelli, F. Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. *J. Hydrometeorol.* **2012**, *13*, 1268–1284. [CrossRef]
6. Cohen Liechti, T.; Matos, G.D.S.C.; Pedro, J.; Boillat, J.-L.; Schleiss, A. 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]
7. Tong, K.; Su, F.; Yang, D.; Hao, Z. Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *J. Hydrol.* **2014**, *519*, 423–437. [CrossRef]
8. 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.* **2016**, *61*, 420–440. [CrossRef]
9. Cheema, M.J.M.; Bastiaanssen, W.G. Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *Int. J. Remote Sens.* **2012**, *33*, 2603–2627. [CrossRef]
10. Iqbal, M.F.; Athar, H. Validation of satellite based precipitation over diverse topography of Pakistan. *Atmos. Res.* **2018**, *201*, 247–260. [CrossRef]
11. 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]

12. Raich, R.; Alpert, P.; Messer, H. Vertical precipitation estimation using microwave links in conjunction with weather radar. In Proceedings of the 15th International Conference on Environmental Science and Technology, Rhodes, Greece, 31 August–2 September 2017.
13. Tapiador, F.; Navarro, A.; Levizzani, V.; García-Ortega, E.; Huffman, G.; Kidd, C.; Kucera, P.; Kummerow, C.; Masunaga, H.; Petersen, W. Global precipitation measurements for validating climate models. *Atmos. Res.* **2017**, *197*, 1–20. [CrossRef]
14. 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]
15. Hartmann, D.; Klein Tank, A.; Rusticucci, M. Observations: Atmosphere and surface, chap. 2 in: Climate change 2013: The physical science basis. In *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013.
16. Munoz, E.A.; Di Paola, F.; Lanfri, M.A. Advances on rain rate retrieval from satellite platforms using artificial neural networks. *IEEE Lat. Am. Trans.* **2015**, *13*, 3179–3186. [CrossRef]
17. 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]
18. Casella, D.; Dietrich, S.; Di Paola, F.; Formenton, M.; Mugnai, A.; Porcù, F.; Sanò, P. PM-GCD-a combined IR-MW satellite technique for frequent retrieval of heavy precipitation. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 231–240. [CrossRef]
19. 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]
20. Zhu, H.; Li, Y.; Huang, Y.; Li, Y.; Hou, C.; Shi, X. Evaluation and hydrological application of satellite-based precipitation datasets in driving hydrological models over the Huifa river basin in Northeast China. *Atmos. Res.* **2018**, *207*, 28–41. [CrossRef]
21. Lu, Y.; Zhang, W.; Wang, K. Evaluation of heavy precipitation simulated by the WRF model using 4D-Var data assimilation with TRMM 3B42 and GPM IMERG over the Huaihe River Basin, China. *Remote Sens.* **2018**, *10*, 646.
22. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]
23. Prakash, S.; Gairola, R. Validation of TRMM-3B42 precipitation product over the tropical Indian Ocean using rain gauge data from the RAMA buoy array. *Theor. Appl. Climatol.* **2014**, *115*, 451–460. [CrossRef]
24. Dinku, T.; Chidzambwa, S.; Ceccato, P.; Connor, S.; Ropelewski, C. Validation of high-resolution satellite rainfall products over complex terrain. *Int. J. Remote Sens.* **2008**, *29*, 4097–4110. [CrossRef]
25. Pierre, C.; Bergametti, G.; Marticorena, B.; Mougin, E.; Lebel, T.; Ali, A. Pluriannual comparisons of satellite-based rainfall products over the Sahelian belt for seasonal vegetation modeling. *J. Geophys. Res. Atmos.* **2011**, *116*. [CrossRef]
26. Gosset, M.; Viarre, J.; Quantin, G.; Alcoba, M. Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Q. J. R. Meteorol. Soc.* **2013**, *139*, 923–940. [CrossRef]
27. Ochoa, A.; Pineda, L.; Crespo, P.; Willems, P. Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific–Andean region of Ecuador and Peru. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 3179–3193. [CrossRef]
28. Salio, P.; Hobouchian, M.P.; Skabar, Y.G.; Vila, D. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmos. Res.* **2015**, *163*, 146–161. [CrossRef]
29. Blacutt, L.A.; Herdies, D.L.; de Gonçalves, 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]
30. Satgé, F.; Bonnet, M.-P.; Gosset, M.; Molina, J.; Lima, W.H.Y.; Zolá, R.P.; Timouk, F.; Garnier, J. Assessment of satellite rainfall products over the Andean plateau. *Atmos. Res.* **2016**, *167*, 1–14. [CrossRef]
31. Kidd, C.; Bauer, P.; Turk, J.; Huffman, G.; Joyce, R.; Hsu, K.-L.; Braithwaite, D. Intercomparison of high-resolution precipitation products over northwest Europe. *J. Hydrometeorol.* **2012**, *13*, 67–83. [CrossRef]

32. Stampoulis, D.; Anagnostou, E.N.; Nikolopoulos, E.I. Assessment of high-resolution satellite-based rainfall estimates over the Mediterranean during heavy precipitation events. *J. Hydrometeorol.* **2013**, *14*, 1500–1514. [CrossRef]
33. Conti, F.L.; Hsu, K.-L.; Noto, L.V.; Sorooshian, S. Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea. *Atmos. Res.* **2014**, *138*, 189–204. [CrossRef]
34. 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]
35. Chen, Y.; Ebert, E.E.; Walsh, K.J.; Davidson, N.E. Evaluation of TMPA 3B42 daily precipitation estimates of tropical cyclone rainfall over Australia. *J. Geophys. Res. Atmos.* **2013**, *118*, 911–966. [CrossRef]
36. Peña-Arancibia, J.L.; van Dijk, A.I.; Renzullo, L.J.; Mulligan, M. Evaluation of precipitation estimation accuracy in reanalyses, satellite products, and an ensemble method for regions in Australia and South and East Asia. *J. Hydrometeorol.* **2013**, *14*, 1323–1333. [CrossRef]
37. Pipunic, R.; Ryu, D.; Costelloe, J.; Su, C. *Evaluation of Real-Time Satellite Rainfall Products in Semi-Arid/Arid Australia*; MODSIM2013; Piantadosi, J., Anderssen, R.S., Boland, J., Eds.; Modelling and Simulation Society of Australia & New Zealand: Canberra, Australia, 2013; p. 31063112.
38. 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]
39. 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]
40. Yuan, F.; Zhang, L.; Win, K.W.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Assessment of GPM and TRMM multi-satellite precipitation products in streamflow simulations in a data-sparse mountainous watershed in Myanmar. *Remote Sens.* **2017**, *9*, 302. [CrossRef]
41. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.; 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. [CrossRef]
42. Muhammad, W.; Yang, H.; Lei, H.; Muhammad, A.; Yang, D. Improving the regional applicability of satellite precipitation products by ensemble algorithm. *Remote Sens.* **2018**, *10*, 577. [CrossRef]
43. 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.* **2017**, *131*, 1–14. [CrossRef]
44. Anjum, M.N.; Ding, Y.; Shangguan, D.; Ijaz, M.W.; Zhang, S. Evaluation of high-resolution satellite-based real-time and post-real-time precipitation estimates during 2010 extreme flood event in Swat River Basin, Hindukush region. *Adv. Meteorol.* **2016**, *2016*. [CrossRef]
45. Butt, M.J.; Iqbal, M.F. Impact of climate variability on snow cover: A case study of northern Pakistan. *Pak. J. Meteorol.* **2009**, *5*, 10.
46. 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]
47. 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*, 1–16. [CrossRef]
48. Asmat, U.; Athar, H.; Nabeel, A.; Latif, M. An AOGCM based assessment of interseasonal variability in Pakistan. *Clim. Dyn.* **2017**, *50*, 1–25. [CrossRef]
49. Salma, S.; Shah, M.; Rehman, S. Rainfall trends in different climate zones of Pakistan. *Pak. J. Meteorol.* **2012**, *9*, 17.
50. Greene, W.H. *Econometric Analysis*; Pearson Education India: Delhi, India, 2003.
51. Gujarati, D.N. *Basic Econometrics*; Tata McGraw-Hill Education: Noida, India, 2009.
52. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Day 1 IMERG Final Run Release Notes*; NASA/GSFC: Greenbelt, MD, USA, 2015.
53. Simpson, J.; Adler, R.F.; North, G.R. A proposed tropical rainfall measuring mission (TRMM) satellite. *Bull. Am. Meteorol. Soc.* **1988**, *69*, 278–295. [CrossRef]

54. 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]
55. 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]
56. As-Syakur, A.; Tanaka, T.; Prasetia, R.; Swardika, I.; Kasa, I. Comparison of TRMM multisatellite precipitation analysis (TMPA) products and daily-monthly gauge data over Bali. *Int. J. Remote Sens.* **2011**, *32*, 8969–8982. [CrossRef]
57. 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]
58. Gebregiorgis, A.S.; Hossain, F. Understanding the dependence of satellite rainfall uncertainty on topography and climate for hydrologic model simulation. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 704–718. [CrossRef]
59. Hirpa, F.A.; Gebremichael, M.; Hopson, T. Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 1044–1051. [CrossRef]

remote sensing

MDPI

Article

Comparison of the GPM IMERG Final Precipitation Product to RADOLAN Weather Radar Data over the Topographically and Climatically Diverse Germany

Thomas Ramsauer *, Thomas Weiß and Philip Marzahn

Department of Geography, LMU Munich, Luisenstraße 37, 80333 Munich, Germany;
t.weiss@iggf.geo.uni-muenchen.de (T.W.); p.marzahn@iggf.geo.uni-muenchen.de (P.M.)
* Correspondence: t.ramsauer@iggf.geo.uni-muenchen.de; Tel.: +49-(0)89-2180-4183

Received: 15 October 2018; Accepted: 2 December 2018; Published: 13 December 2018

Abstract: Precipitation measurements provide crucial information for hydrometeorological applications. In regions where typical precipitation measurement gauges are sparse, gridded products aim to provide alternative data sources. This study examines the performance of NASA's Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement Mission (IMERG, GPM) satellite precipitation dataset in capturing the spatio-temporal variability of weather events compared to the German weather radar dataset RADOLAN RW. Besides quantity, also timing of rainfall is of very high importance when modeling or monitoring the hydrologic cycle. Therefore, detection metrics are evaluated along with standard statistical measures to test both datasets. Using indices like "probability of detection" allows a binary evaluation showing the basic categorical accordance of the radar and satellite data. Furthermore, a pixel-by-pixel comparison is performed to assess the ability to represent the spatial variability of rainfall and precipitation quantity. All calculations are additionally carried out for seasonal subsets of the data to assess potentially different behavior due to differences in precipitation schemes. The results indicate significant differences between the datasets. Overall, GPM IMERG overestimates the quantity of precipitation compared to RADOLAN, especially in the winter season. Moreover, shortcomings in detection performance arise in this season with significant erroneously-detected, yet also missed precipitation events compared to the weather radar data. Additionally, along secondary mountain ranges and the Alps, topographically-induced precipitation is not represented in GPM data, which generally shows a lack of spatial variability in rainfall and snowfall estimates due to lower resolution.

Keywords: precipitation; weather; radar; GPM; RADOLAN; QPE

1. Introduction

Precipitation is of paramount importance as a driver of the global water and energy cycle and interactions between the bio-, hydro-, and atmosphere and thus has been declared as an Essential Climate Variable (ECV) [1]. Information on the spatial and temporal distribution of this crucial variable helps in understanding its vast impact on numerous environmental aspects of life on Earth. Water resource management, predicting and monitoring agricultural yields, or disaster prevention and ultimately management are exemplary fields that strongly depend on accurate precipitation measurements. Traditional measurement gauges are sparse in many parts of the world [2], which hindered the deduction of meaningful precipitation estimates for these regions until a few decades ago, when gridded (satellite) products came to close these gaps. Currently, a physically-measured precipitation distribution can be acquired via interpolation of gauge measurements, weather radar estimates, or satellite observation. At the global scale, the spatial variability of rain and snowfall can be best represented with remote sensing imagery, as radar and gauge measurement stations are not

available world-wide with sufficient density and coverage. Moreover, time-series of satellite data let global precipitation patterns and distribution become apparent. Still, region-specific differences in climate and topography are determinant factors for uncertainties in the performance of satellite precipitation products. Currently, developments to improve gridded precipitation data utilize creation or correction approaches for satellite-based precipitation products (SPP) from satellite soil moisture retrieval data [3–8] or combine datasets from various sources like gauge measurements, atmospheric models, and satellite observations [9].

NASA's Global Precipitation Measurement (GPM) mission launched the GPM Core Observatory (CO) as the successor of the well-renowned Tropical Rainfall Measuring Mission (TRMM) spacecraft in 2014 [10]. Additional channels on both the Dual-frequency Precipitation Radar (DPR) and on the GPM Microwave Imager (GMI) make it an advanced replacement of the older satellite. The Integrated Multi-Satellite Retrievals for GPM (IMERG) gridded dataset used in this study is a Level 3 NASA product which, unifies and inter-calibrates data of about 10 constellation satellites from several space agencies based on the GPM CO [11–13].

Numerous comparison studies involving GPM data have been carried out over different spatial domains, e.g., global [14], Canada [15], Singapore [16], Malaysia [17], China [18–20], India [21], Iran [22], and Saudi-Arabia [23]; yet, investigations covering European countries are sparse, and no detailed comparison over Germany exists until today. However, the consistent availability at high temporal and spatial resolution and hence lowered uncertainty propagation in the results of hydrological modeling make GPM a viable data source for applications across European catchments of different scales [24]. Nevertheless, systematic bias and random errors are usually contained in satellite precipitation estimates [25,26]. Mei et al. [27] showed that SPPs furthermore are prone to underestimation of extreme events and hence are the main contributor to the total error in their hydrological modeling setup. Although GPM data are currently barely used in hydrology-related modeling scenarios in Europe, numerous future applications have been proposed. The topics cover, e.g., landslide threshold precipitation in the Italian Umbria region [28], debris flow-triggering rainfall [29], or modeling of flood events in alpine terrain [30]. Moreover, GPM data are now incorporated in the Global Flood Detection System (GFDS [31]) [32]. The insufficient performance of this dataset over Germany, which has been demonstrated in a validation study in the TRMM era [33], generates uncertainty for future usage. Hence, a performance test of GPM over Germany is necessary, to allow questioning these kinds of results over this or similar geographic regions.

Furthermore, the existing comparison setups include different datasets. Speirs et al. [34] for example compared GPM DPR to the MeteoSwiss radar network with a focus on mountainous regions. The radar data are adjusted, yet only to a very limited number of gauges (6–10, 33) and not on an operational basis, but to long-term mean precipitation values. Other studies also evaluated GPM (and mostly the DPR product) against weather radar datasets [35–37] where many focused on performance towards snow detection [38–43]. The resulting findings indicate huge improvements compared to the TRMM era. Yet, the need for future improvements of the algorithm to further enhance the IMERG abilities in freezing conditions still persists [22,34,40,44,45].

Studies on the performance of SPPs are strongly location dependent with highly diverse correlation values to gauge measurements especially in challenging topography [17,46]. Therefore, the evaluation of quantitative precipitation estimates (QPE) is vital before operationally applying them in a specific study site. Germany, in additional to its diverse topography, lies in the transition zone from oceanic to continental climate with different apparent precipitation schemes, making it a very interesting and challenging case study.

The novelty in the presented case is the comparison of the final GPM IMERG data to a temporal and spatial high resolution precipitation product. This product is the state-of-the-art weather radar-derived and operationally gauge-adjusted precipitation product RADOLAN RW from the German Weather Service (DWD, Deutscher Wetter Dienst). Due to the high sampling frequency, short-scale precipitation events can be captured. Furthermore, the hourly online adjustment routine

makes it a balanced dataset, adhering to a high degree to the gauge measurements without cutting out extreme events [47,48].

To assess the performance of GPM over complex terrain, throughout seasons and consequently on different precipitation regimes, the study aims to compare final GPM IMERG against RADOLAN RW data from DWD. Therefore, different standard statistical measures, as well as a range of categorical indices are applied and evaluated on a pixel-by-pixel basis. Utilizing this form of spatial comparison accounts for the drastic topographic differences throughout the study area with landscapes including lowlands, secondary mountain ranges to alpine peaks with heights up to 3000 m.a.s.l., as well as for the different seasons and precipitation regimes. Thus, two hypotheses will be addressed throughout the study: (1) GPM shows similar detection performance over different topographic and climate zones compared to RADOLAN data; (2) GPM and RADOLAN show the same spatial and seasonal trends in precipitation.

2. Study Area

The spatial bounds for the dataset comparison are comprised of the state territory of Germany, which extends from 47° to 55°N and from 5° to 16°E, respectively, and covers an area of 357,021 km^2.

The topography is diverse, with lowlands in the north, uplands and secondary mountain ranges in central region and the foothills of the Alps, and adjacent summits with their highest peak being Zugspitze (2962 m.a.s.l.) in the southern part of Germany. An overview of the study area is given in Figure 1. Accordingly, the relief variability increases towards the southern part, where strong gradients in temperature and precipitation are caused by steep slopes in the mountainous region over a very short horizontal distance. For example, Garmisch-Partenkirchen at the foot of Zugspitze is characterized by a mean temperature of 7.2 °C and annual precipitation of 1231 mm, whereas the summit weather station yields −3.7 °C and 1978 mm. Overall, a temperate seasonal climate prevails with mean temperatures ranging from −3.7 °C to 11.0 °C and a mean annual precipitation ranging from 483 mm to 2340 mm.

The distribution of precipitation in Germany is induced by the spatial position of the state lying in between the oceanic Western Europe and the continental Eastern Europe. Amounts of precipitation, mostly brought by humid westerly winds, decrease towards the eastern parts of the study area, yet regions in the extreme south and parts of the uplands in central Germany show higher precipitation amounts due to their mountainous climate. In the winter time, solid precipitation in the form of snow is more common in areas with continental influence.

The time period from 1 December 2014 to 30 November 2017 is analyzed in this study.

3. Data and Methodology

3.1. Datasets

3.1.1. Weather Radar Data

The gauge-adjusted quality-controlled RADOLAN RW (Radar Online Adjustment) dataset from the German Weather Service (DWD, Deutscher Wetter Dienst) is considered ground truth for the upcoming analyses. It is already widely used, e.g., for training and validation purposes in the machine learning domain [49,50], analyzing extreme flash floods [51], as well as enhancing the respective forecasts [52] and estimating the spatio-temporal variability of soil erosion [53].

The radar dataset is currently derived from 18 C-band weather radars operating on scanning intervals of 5 minutes. All but the radar station “Hohenpeißenberg”, which is used for quality control, contribute to the quantitative precipitation analysis. The observational network’s spatial distribution is shown in Figure 1 along with the associated coverage of each device with a radius of 150 km. Significant overlap within the dense radar network ensures accurate retrievals, since problems from dampening in the signal with increased distance from the sensor and hence missing or misinterpreting

precipitation events are minimized [54]. In the last few years, the weather radars have been gradually updated to dual-polarized scanning devices that allow discriminating the sort of hydrometeors [54]. Within the specific calibration procedure, rain intensity-adapted Z-R relationships (empirical formula to estimate rainfall rates from reflectivity signal strength) and statistical clutter filtering are applied, and orographic shadowing effects are considered [48,55,56].

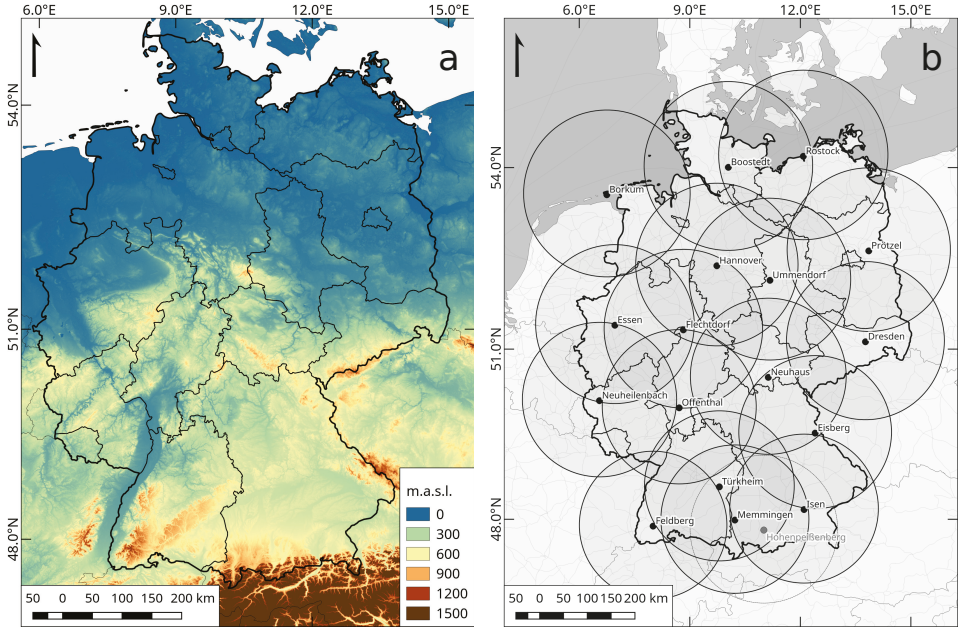


Figure 1. Digital elevation model of Germany based on SRTM 1 arc second data (**a**) and the observational network of weather radar stations contributing to the RADOLAN dataset (**b**).

Assumptions on the drop size distribution and droplet count are necessary for the deduction of precipitation [54]. For RADOLAN, an extended Z-R relationship is utilized, as opposed to solely using standardized values from the literature. The relationship takes the absolute reflectivity, as well as horizontal gradients into account to distinguish typical convective and stratiform droplet distributions [48]. Furthermore, potential overshooting effects in wintertime due to lower cloud heights are considered with a seasonally-dependent correction via a regression analysis. However, a general linear correction scheme does not fulfill the requirements of DWD due to erroneous adaptation of single extreme events, e.g., intensive convective cells that occur regularly throughout Germany in the summer. Therefore, a multiple polynomial regression is calculated to generate the correction factors for every pixel. This accounts for the respective scanning height class, day of year, and reflectivity [54]. The enhancements concerning dual-polarization radar relevant Z-R relationships were not integrated in the online adjustment routine at the time of data acquisition.

Nevertheless, for a realistic estimation of the quantity of precipitation, measurements of approximately 1300 conventional stations are used for the operational hourly gauge adjustment routine [55]. These sensors (Ott PLUVIO) basically work according to "Hellmann" ombrometers [57], which obey the standards of the World Meteorological Organization [58]. The appliance of a weighing principle and surrounding temperature-dependent heating sets the PLUVIO apart from conventional measurement systems and allows capturing solid and fluid precipitation alike [48]. A subset of the

gauge stations is used in the generation of the monthly Global Precipitation Climatology Centre (GPCC) product.

The precipitation product is available at a temporal, spatial, and intensity resolution of 1 h, 1 km, and 0.1 mm. A dimension of 900 × 900 pixels allows the polar-stereographic composite grid with the center point at 9.0°E 51.0°N to cover the whole state territory of Germany [47,48]. Throughout this study, the dataset will be referred to as "RADOLAN".

3.1.2. Satellite Data

The GPM IMERG Version 5 final precipitation half hourly dataset with 0.1-degree spatial resolution is compared to the aforementioned radar precipitation dataset. The GPM Core Satellite is equipped with a multi-channel, dual-polarization Passive Microwave sensor (PMW) and an active scanning radar. Improvements to the predecessor TRMM satellite include increased orbital inclination from 35° to 65° for improved coverage, upgraded radar to two frequencies, as well as additional "high-frequency" channels in the PMW, both allowing for and facilitating the detection of light and solid precipitation, respectively [12,13]. In Version 5, the research-level "final" dataset is adjusted monthly to the extensive GPCC gauge-based dataset, which is available at 1.0° × 1.0° spatial resolution [59]. In the study, the dataset will be addressed as "GPM".

3.1.3. Preprocessing of Datasets

In order to make the datasets spatially and temporally comparable, the RADOLAN dataset was reprojected from the DWD-specific stereographic projection to WGS84, remapped, and aggregated to the GPM grid. Remapping routines using bilinear interpolation or high-order finite-differencing techniques may lead to unexpected behavior, e.g., higher local maxima, and are non-conservative; hence, they behave inconsistent with regard to precipitation sums in the original and regridded dataset [60,61]. Furthermore, bilinear remapping schemes produce significant changes especially to categorical skill scores [62]. Therefore, the ideal regridding scheme to use for precipitation data, being discontinuous over space and time, is the area conservative regridding, which calculates fractional contributions of grid cells from the original data and hence maintains the same area-averaged rainfall before and after the remapping [63]. Thus, the specifically-applied spatial averaging procedure to remap the finer RADOLAN grid data to the coarser GPM grid utilizes the first order conservative remapping scheme from Jones [64], comprised in the Climate Data Operators software (CDO), which applies the SCRIP algorithm (Spherical Coordinate Remapping and Interpolation Package) [65,66]. This technique is widely applied in other studies dealing with precipitation data [67–69]. The area-averaged precipitation quantity $\bar{F}_k$ at the destination grid is calculated as follows:

$$\bar{F}_k = \frac{1}{A_k} \int_{A_k} f dA \tag{1}$$

where A_k denotes the area of the destination grid cell k and f is the precipitation quantity in the original grid, which has an overlapping area with the destination grid [64].

Furthermore, the GPM data were aggregated temporally to match RADOLAN's hourly resolution. Both datasets were clipped to the extent of the state territory of Germany.

3.2. Methodology

The GPM satellite precipitation dataset was statistically compared to RADOLAN weather radar data. Generally, in investigations like this, quality checks of the involved data are critical to produce meaningful results in the end. In this study, 55 weather radar hourly grids are reported as missing, representing solely 0.17% of the considered time steps. The GPM time series is complete. Furthermore, visual interpretation of the radar images for the time span under review indicates no erroneous data

concerning typical radar-related errors like beam blockage and artifacts, which occurred in the first versions of the distributed RADOLAN data at the beginning of the recording period.

To determine whether the datasets show seasonally-dependent dissimilar behavior, due to different precipitation schemes and the higher prevalence of snowfall in winter, the statistical analysis was split into the four meteorological seasons winter (DJF), spring (MAM), summer (JJA), and fall (SON). Overall, statistical comparisons of precipitation sums and means have been carried out. Pixel-by-pixel difference and correlation analyses were conducted additionally to provide a spatial representation of the level of compliance of the RADOLAN and GPM datasets. Pearson's r was used as the correlation measure.

$$r = \frac{cov(P_{\mathrm{GPM}}, P_{\mathrm{RADOLAN}})}{\sigma_{P_{\mathrm{GPM}}} \sigma_{P_{\mathrm{RADOLAN}}}} \tag{2}$$

Furthermore, the overall unconditional bias B was calculated for the data with the following formula.

$$B = \frac{\sum_{i=1}^{N} P_{\mathrm{GPM}_i}}{\sum_{i=1}^{N} P_{\mathrm{RADOLAN}_i}} \tag{3}$$

A perfect linearity of precipitation measurement amounts in GPM and RADOLAN results in a value of 1.

To represent the average magnitude of the error, the Mean Absolute Error (MAE) is used:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N} |P_{\mathrm{GPM}_i} - P_{\mathrm{RADOLAN}_i}|}{N} \tag{4}$$

The Root Mean Squared Error (RMSE) with greater weight for larger errors than the aforementioned MAE is also part of the statistical evaluation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_{\mathrm{GPM}_i} - P_{\mathrm{RADOLAN}_i})^2} \tag{5}$$

where P_{GPM} and P_{RADOLAN} are the satellite and weather radar precipitation estimates, respectively, i denotes the ith hourly event in the case of the pixel-by-pixel calculation, and the ith element (all pixels over all time steps) for the overall calculation. In the same way, N stands for observed hourly values per pixel or the product of the count of pixels and the count of hourly values, respectively.

Furthermore, the ability to ascertain wet days with precipitation amounts greater than 1 mm was examined to allow for inferences to be made about the detection rates of the two precipitation datasets. Therefore, the count of these days and the respective mean precipitation sum have been evaluated for the datasets on a seasonal basis.

Additionally, categorical indices are calculated to further the knowledge about detection performance. They allow the evaluation of the binary accordance of the precipitation datasets, meaning to see if events are captured uniformly in both datasets. This has been done for the spatially-aggregated datasets, as well as on a pixel-by-pixel basis. For these calculations, the contingency grid shown in Table 1 is used, where a, b, c, and d represent the total count of data pairs matching the requested criteria. RADOLAN is chosen as reference due to the originally higher spatial resolution and the higher temporal frequency in adjusting to gauge measurements. For further information on the metrics used, please refer to, e.g., Woodcock [70], Doswell et al. [71], Schaefer [72].

Table 1. Contingency table for the calculation of categorical indices.

		RADOLAN	
		Rain	**No Rain**
GPM	**rain**	a	b
	no rain	c	d

The Probability Of Detection (POD) for GPM measurements over Germany in the reported time period can be written as:

$$\mathrm{POD} = \frac{a}{(a+c)}, \tag{6}$$

and gives a measure of how effective the satellite observations detect a rain event compared to RADOLAN with the perfect score being 1.

The opposite case, where precipitation is erroneously indicated by GPM, is assessed with the False Alarm Ratio (FAR):

$$\mathrm{FAR} = \frac{b}{(a+b)}, \tag{7}$$

where the perfect score is 0.

The Frequency Bias Index (FBI) is the ratio of the total count of precipitation events of the two datasets. The values range from 0 to ∞, with a perfect score of 1:

$$\mathrm{FBI} = \frac{(a+b)}{(a+c)}. \tag{8}$$

This complements the similar measure of the unconditional bias in that the amounts of precipitation are left out and only temporal and spatial similarities in the occurrence of such events are taken into consideration.

The Critical Success Index (CSI) combines the information of FAR and POD. Thus, it shows how well the correctly-detected precipitation events from GPM conform to all the recorded precipitation events, making the CSI a very balanced measure, with the best score being 1:

$$\mathrm{CSI} = \frac{a}{(a+b+c)}. \tag{9}$$

Finally, the Heidke Skill Score (HSS) was calculated for the datasets. This metric answers the question on accuracy against random guessing. For a perfect measurement, the value will be 1. Performance equal to or worse than random guessing results in $-1 \leq \mathrm{HSS} \leq 0$:

$$\mathrm{HSS} = \frac{2 \times (a \times d - b \times c)}{((a+c) \times (c+d) + (a+b) \times (b+d))}. \tag{10}$$

A threshold of 0.1 mm/h is defined to delineate a precipitation event for the calculation of the above indices. This is in agreement with both datasets' intensity resolution. Hourly pixel values below this threshold are treated as noise and therefore are omitted.

4. Results

4.1. Statistical Analysis

4.1.1. Overall

Figure 2a,b shows the yearly mean precipitation of the two datasets. The plots serve clearly as evidence for the different recording techniques and their initially different spatial properties. The topographic characteristics of Germany can be traced from the RADOLAN data, which, although

spatially aggregated, reveal the inherited higher spatial variability. In contrast, the yearly mean precipitation measured by the GPM constellation appears smoother. The overall pattern indicates a similar precipitation distribution across Germany with high divergence in the level of detail. Both datasets agreed on the foothills of the Alps as the rain-laden region and eastern Germany as the driest sub-region in the state territory. The difference of GPM's and RADOLAN's precipitation amounts over the whole period under review again demonstrates the differences in spatial variability of the datasets. Furthermore, GPM in many parts of Germany overestimated the quantity of precipitation. Yet, over areas of secondary mountain ranges and alpine regions, the satellite data indicated lower precipitation amounts than the gauge-adjusted weather radar (Figure 2c).

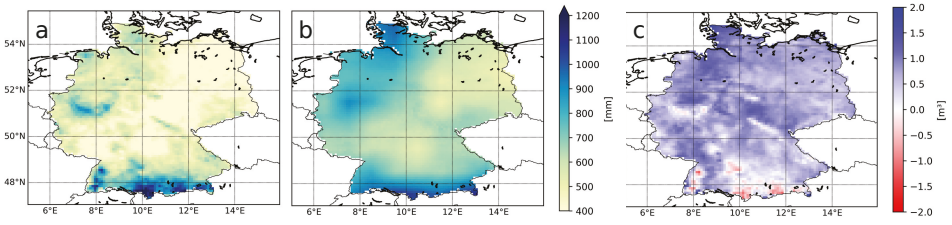


Figure 2. Yearly mean precipitation sum over Germany from RADOLAN (**a**) and GPM (**b**) data and the overall difference (GPM − RADOLAN) calculated for the period under review (**c**).

The monthly precipitation sums averaged over entire Germany show a clear pattern (Figure 3). Across all winter months in the reporting period, GPM's QPE clearly exceeded those of RADOLAN with a maximum monthly mean surplus per pixel of >20 mm. In summer months, the collected data coincided.

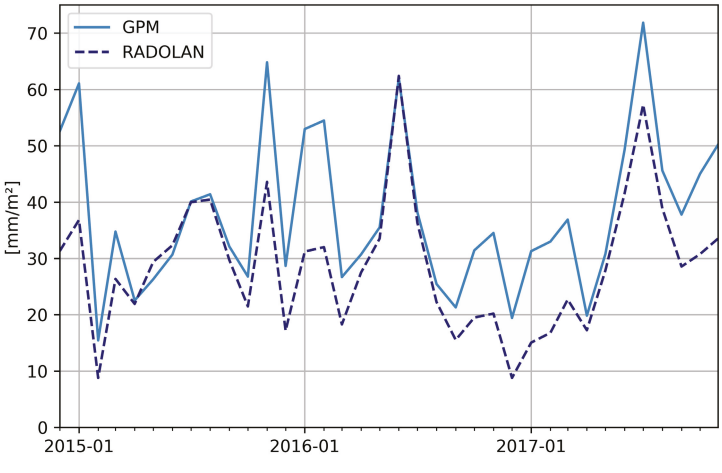


Figure 3. Spatially-averaged monthly precipitation sums in the GPM and RADOLAN datasets.

The evaluation of the unconditional bias upholds previous findings by also indicating a general overestimation of the precipitation amount by the GPM data compared to RADOLAN's QPE with $B = 1.31$.

4.1.2. Seasonal Analysis

The analysis of seasonal aggregated data was used to further reveal differences in precipitation patterns and the respective detection by the GPM and RADOLAN datasets. The absolute differences per season over the whole reporting period are shown in Figure 4. Besides the again prominent existence of differences due to spatial variability, the differences are diverse across seasons and conform to Figure 3. In fall and winter months (SON, DJF; Figure 4), GPM data showed higher precipitation values than RADOLAN in most areas. In the other two seasons, the satellite QPE were generally more on par with the weather radar data. However, in the southern part of the study area, RADOLAN showed higher values in spring (MAM) and especially in the summer season (JJA).

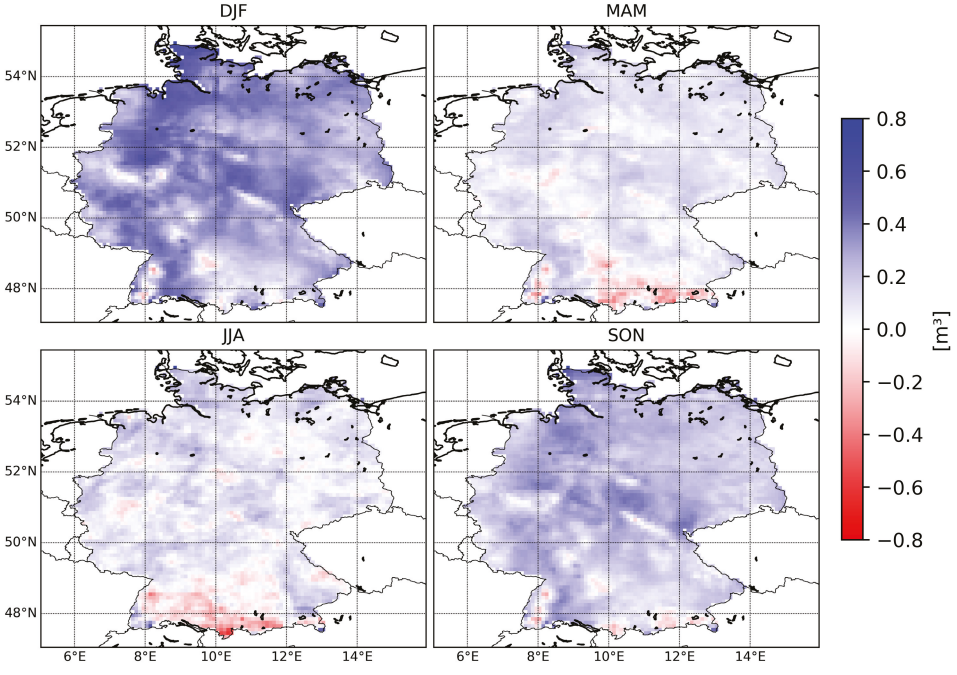


Figure 4. Differences in precipitation sums from GPM-RADOLAN datasets for seasons DJF, MAM, JJA, and SON over the reporting period.

These findings are further supported by the mean precipitation sums per season across the territory of Germany, which are shown in Figure 5a. An overestimation of the precipitation amount by GPM data occurred in all seasons. However, wintertime with a surplus of 76% needs to be emphasized.

Pearson's R value was utilized to calculate the correlation between GPM and RADOLAN precipitation. Additionally, the measure was applied on a pixel-by-pixel basis to evaluate the GPM and RADOLAN data's spatial agreement. Therefore, for every location in every seasonal data subset, the correlation was calculated. The overall correlation was 0.49, where for the single seasons, the values differed greatly, resulting in a value of 0.38 for DJF-, 0.55 for MAM-, 0.54 for JJA-, and 0.57 for SON-season. These results were backed by the spatial representations shown in Figure 6. All seasons besides DJF showed moderate correlation throughout the state territory of Germany. In the winter season, however, great shares of the southeastern parts of the study region showed very low correlation values around 0.1 to 0.2 with the minimum being 0.07.

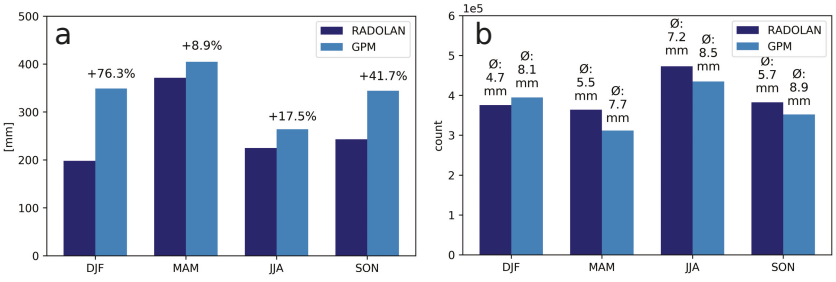


Figure 5. Mean of seasonal precipitation sums (**a**) and seasonal count and mean precipitation of "wet days" (**b**) of the GPM and RADOLAN datasets.

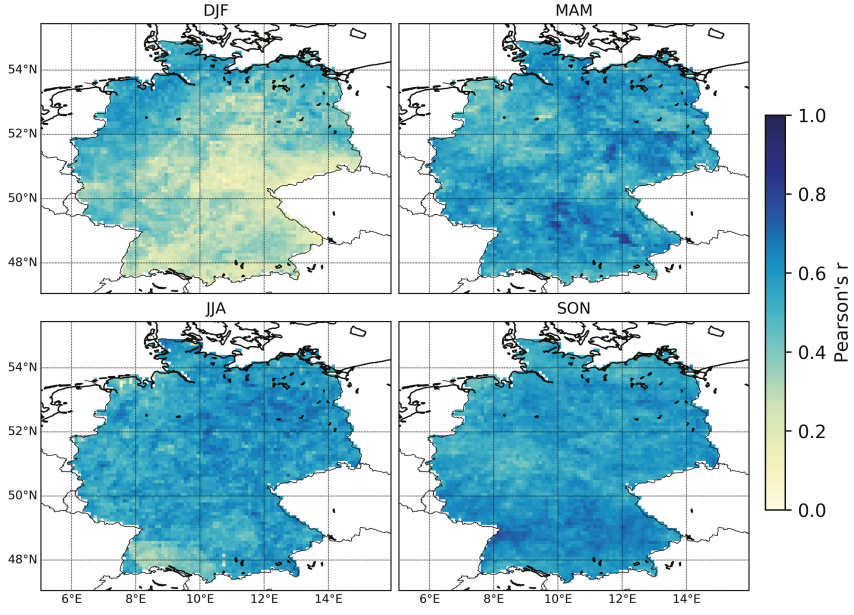


Figure 6. Pixel-by-pixel correlation of precipitation from GPM and RADOLAN datasets for seasons DJF, MAM, JJA, and SON.

4.2. Categorical Performance

The amount of wet days with a daily precipitation sum greater than 1 mm varied between the datasets (see Figure 5b). Besides the winter season, RADOLAN captured significantly more rain events than GPM. In the spring season, this accounted for up to 50,000 pixel hours within the reporting period. Yet, also in accordance with previous results, the mean precipitation amount per wet day measured by GPM was higher than the respective RADOLAN value in all seasons. Although GPM showed a lower detection rate for wet days, the surplus of precipitation amount compensated this effect, allowing the aforementioned results concerning the satellite measurements to be positively biased compared to RADOLAN to still be valid.

Diverse categorical indices have been calculated to obtain knowledge about the dataset-specific detection capabilities concerning precipitation events (see Section 3.2). These were again calculated for

the whole datasets, as well as for seasonal subsets. Furthermore, a spatial representation calculated on a pixel basis may be found in Figures 7 and 8.

The capability of the GPM dataset to capture every precipitation event was moderate with an overall value of 0.53 (see Table 2). Regions with high relief energy showed the lowest POD values throughout all the seasons. The highest amounts of erroneously-detected precipitation events showed up in the eastern part of Germany, demarcated most clearly in the SON and DJF seasons. This demarcation is related directly to the FBI being strongly positive in that region in the same seasons. Still, more events per pixels across Germany were detected by RADOLAN in all seasons, resulting in values of FBI ranging from 0.68 to 0.90 with an overall value of 0.78.

A different temporal pattern can be found in the error indices MAE and RMSE. However, due to the aforementioned topography related concern, the spatial shortcomings of GPM versus RADOLAN in representing precipitation still persisted. Besides the winter season, also in the summer, high error values throughout most of Germany were present. Nevertheless, alpine regions have to be highlighted as specific region, as the error values clearly exceeded the error values from the rest of Germany.

Table 2. Categorical indices per season.

Season	All	DJF	MAM	JJA	SON
Probability Of Detection (POD)	0.53	0.38	0.51	0.67	0.56
False Alarm Ratio (FAR)	0.32	0.48	0.25	0.25	0.32
Critical Success Index (CSI)	0.42	0.28	0.43	0.55	0.44
Heidke Skill Score (HSS)	0.56	0.39	0.57	0.68	0.58
Frequency Bias Index (FBI)	0.78	0.72	0.68	0.90	0.83
Mean Absolute Error (MAE)	0.11	0.12	0.08	0.12	0.10
Root Mean Square Error (RMSE)	0.59	0.66	0.47	0.67	0.52

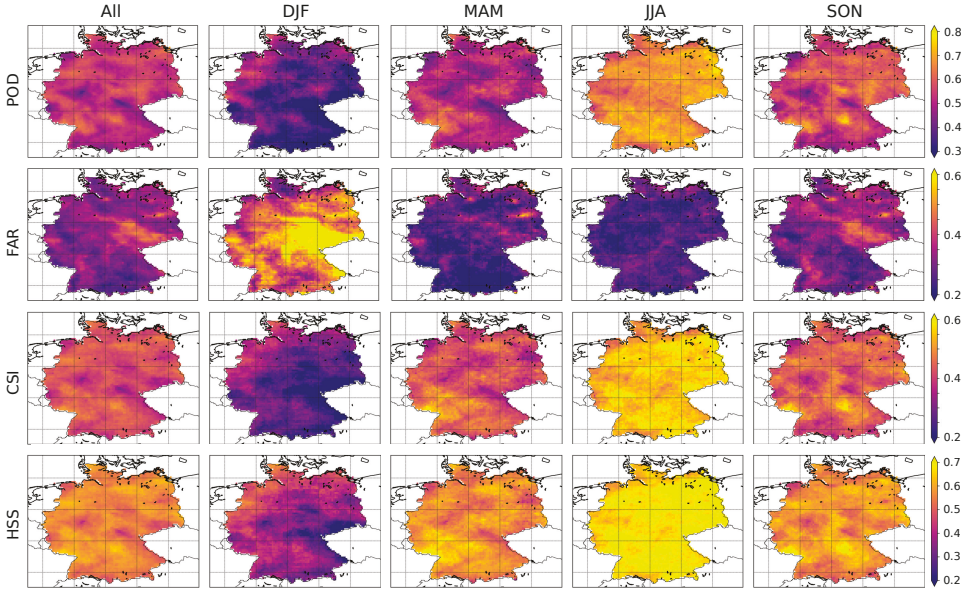


Figure 7. Categorical indices POD, FAR, CSI, and HSS for the total review period, DJF, MAM, JJA, and SON seasons.

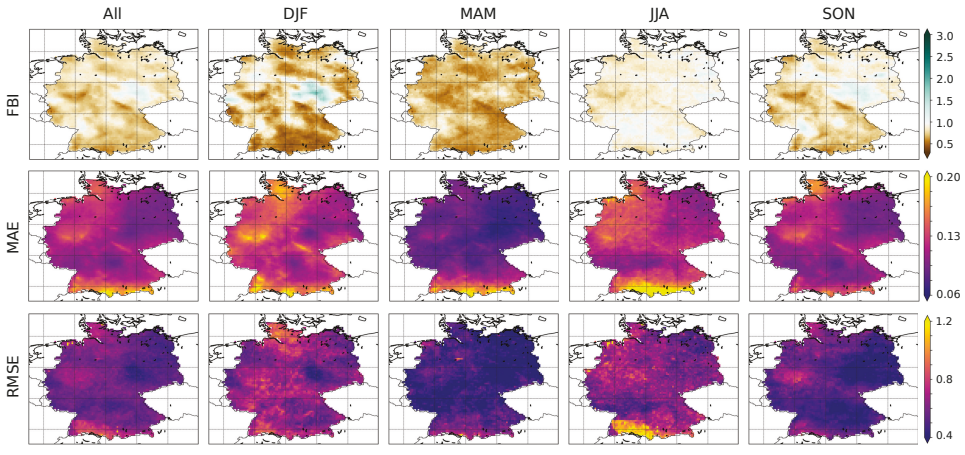


Figure 8. FBI, MAE, and RMSE for for total review period, DJF, MAM, JJA, and SON seasons.

5. Discussion

The single most marked observation to emerge from the data comparison is the strong discrepancy of the GPM and RADOLAN dataset concerning precipitation estimation for the winter season. Correlation between the satellite observation and weather radar data is low for this time period and seems to show an inversely proportional relation to continentality. Combined with low POD values, uncertainty arises with respect to the applicability of the dataset in, e.g., hydrological modeling. The problems of GPM dealing with solid precipitation have to be considered as one rationale behind the low detection rate, yet highly overestimated precipitation amounts in the winter season compared to the weather radar. GPM IMERG data being positively and negatively biased in cold environments is consistent with previous findings in the literature reported from [22,36], respectively. He et al. [18] even excluded winter months from their study as both satellite and gauge measurements are error prone in the detection of solid precipitation. Kochendorfer et al. [73] also stated that weighing precipitation gauges is highly error prone, especially when wind speeds exceed 5 ms^{-1}. In this case, less than 50% of the actual amount of solid precipitation may be collected. For the type of measurement gauges mainly used in Germany, Boudala et al. [74] reported an undercatch with a ratio of 0.57 for solid precipitation. Different filter algorithms are applied to the gauge measurements by the DWD. However, wind effects may still alter the measurements [48].

In the current study, GPM was positively biased compared with RADOLAN throughout all seasons. Biased precipitation estimation of the satellite dataset has been published by several authors [17,18,22,44], however, for both positive and negative directions. Furthermore, the already mentioned results of quantitative overestimation, particularly in winter and partly caused by false alarms, account for the shift in the precipitation amounts. The very high FAR and FBI values in eastern Germany in winter (see Figures 7 and 8), where lakes and big rivers (Elbe, Havel, Mulde) are abundant lead to an assumption of these landscapes and their inherent water cycle influencing the retrieval. Although, there is a high discrepancy in the number of events, there is no sign of excessive overestimation of the quantity of precipitation compared to the surrounding regions. Thus, e.g., ground fog, possibly not detected by RADOLAN though overrepresented in GPM, could be taken into consideration as an explanation for the disagreement of both datasets. Furthermore, solid precipitation in winter could be the reason for the discrepancies, although other areas throughout Germany are definitely more prone to snowfall. Moreover, the region is located in the lee of a secondary mountain range. Erroneously-detected precipitation in areas of rain shadow is reported for GPM estimates by Prakash et al. [21]. The performance of GPM considering light rain and solid particle detection

increased compared to its predecessor TRMM [11], yet the present case demonstrates like other studies that the need to further improve the algorithm still exists.

Furthermore, the detection of orographic precipitation is erroneous in GPM, which has already been covered by several studies [18,21,22]. The inability to capture topography-induced convective precipitation clearly becomes evident in this study by most categorical indices and the overall difference image signifying these areas (see Figures 2c, 7 and 8). Therefore, existing high rainfall intensities along the Alps naturally lead to high error values in MAE and RMSE. The grainy nature of RMSE in the summer season JJA (see Figure 8) and high error values in southwest Germany can certainly be attributed to the nature of the metric itself and hence to the sensitivity towards high intensity precipitation events, which commonly occur in these regions throughout summer. Due to the RADOLAN's inherent shorter scanning interval and thus, after aggregation to hourly data, still existing higher probability to detect a high intensity rainfall happening on a short temporal scale, great discrepancies in the RMSE may arise from a missed precipitation event by GPM, particularly in the summer season.

General caution has to be applied when datasets with originally different spatial resolution are compared. Although the applied conservative remapping scheme as described in Section 3.1.3 is widely appreciated as very suitable for regridding precipitation data, other techniques (e.g., bilinear, bicubic, iterative curvature-based interpolation) may slightly alter the findings of this study. However, the authors compared the results from the highly unequal non-conservative bilinear interpolation (data not shown) and the applied conservative interpolation scheme, finding that the changes in the results were very small and did not change the statement of the results. However, we recommend that future studies should consider an in depth analysis of the impact of the different interpolation schemes on the comparison of different precipitation datasets. Moreover, a transferability of the results can only be given to regions that share similar boundary conditions. Therefore, the case study over Germany is well suited, as it represents various topographical conditions, as well as several precipitation regimes to test the performance of GPM.

Further processing could include temporal aggregation to and comparison of daily values as precipitation data often are used on this temporal scale as input for other applications, e.g., in hydrologic modeling. It has to be noted that both institutes, NASA and DWD, provide additional products of the respective family (GPM and RADOLAN), which are calculated with a modified algorithm or are based on a subset of sensors. However, the specific purpose of this study was to compare the respective final community-ready precipitation datasets GPM IMERG v05 final and RADOLAN RW, which fully incorporate all data gathered for the respective mission.

Lastly, it has to be noted that the identified performance-related discrepancies profoundly become popular, as the two data sets cannot be considered entirely independent. GPM IMERG utilizes data from the GPCC network on a monthly basis for calibration. Parts of the involved gauges are also used in the hourly online adjustment routine of the RADOLAN dataset. This issue has been accepted by the authors as the calibration for both datasets takes place on a totally different temporal scale.

6. Conclusions

This study conducted a statistical comparison of two QPE products, namely the GPM IMERG half hourly Version 5 final satellite and RADOLAN RW weather radar dataset. Standard metrics like RMSE, MAE, and bias have been applied and categorical indices used to identify the strengths and shortcomings in the ability to detect single precipitation events. Additionally, a pixel-by-pixel analysis of these measures allows drawing conclusions on the spatial distribution of the inherent event identification capabilities of the GPM and RADOLAN datasets.

The results provide considerable insight into the different properties and indicate extensive discrepancies in some parts of the study. Four key findings are revealed by the analysis: (i) the GPM dataset shows low responsivity for the topographically-induced spatial variability of precipitation over Germany compared to the RADOLAN data (see Figure 2); (ii) the precipitation amounts measured by

the satellite product exceed the weather radar data on a territory scale in all seasons, especially in winter (see Figure 3), whereas over spatial subsets with high relief energy, RADOLAN is on par or generates a surplus in precipitation quantity (see Figure 4 and MAE and RMSE in Figure 8); (iii) RADOLAN captures a higher amount of low intensity events (see the high FBI in Figure 8); and (iv) substantial differences in winter season have to be reported, in terms of low correlation (see Figure 6) and high FAR values, yet low POD and CSI/HSS success statistics (see Figure 7). These outcomes lead to the conclusion that caution and awareness of the peculiarities of the dataset have to be applied when using GPM data over Germany and thus also over parts of Europe. However, this protective measure extends to every dataset, which is attributed to being a reference or used in a similar manner.

Author Contributions: Conceptualization, T.R.; Methodology and Formal Analysis, T.R.; Writing—Original Draft Preparation, T.R.; Writing—Review & Editing, T.R., T.W. and P.M.; Visualization, T.R.; Project Administration, P.M.

Funding: The project leading to this application has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 687320. **Acknowledgments:** Thomas Ramsauer is truly thankful for the inspiration, ideas, and support that Alexander Loew has given to him.

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

References

1. GCOS. *The Global Observing System for Climate: Implementation Needs*. WMO Pub GCOS-200; GCOS: Geneva, Switzerland, 2016.
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]
3. 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]
4. 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]
5. 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]
6. 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]
7. 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]
8. 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]
9. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.J.M.; Weedon, G.; Brocca, L.; Pappenberger, F.; Huffman, G.J.; Wood, E.F. Global-scale evaluation of 23 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci. Discuss.* **2017**, *2017*, 1–23. [CrossRef]
10. Simpson, J.; Adler, R.F.; North, G.R. A proposed tropical rainfall measuring mission (TRMM) satellite. *Bull. Am. Meteorol. Soc.* **1988**, *69*, 278–295. [CrossRef]
11. 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]
12. 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 and Basis Document and (ATBD) and Version 5.2*; Technical Report; National Aeronautics and Space Administration: Greenbelt, MD, USA, 2018.

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. Liu, Z. Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results. *J. Hydrometeorol.* **2016**, *17*, 777–790. [CrossRef]
15. 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]
16. Tan, M.; Duan, Z. Assessment of GPM and TRMM Precipitation Products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]
17. 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]
18. 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]
19. 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]
20. 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]
21. Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Liu, Z.; Norouzi, H.; Pai, D. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.* **2018**, *556*, 865–876. [CrossRef]
22. 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]
23. 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]
24. Nikolopoulos, E.I.; Anagnostou, E.N.; Hossain, F.; Gebremichael, M.; Borga, M. Understanding the Scale Relationships of Uncertainty Propagation of Satellite Rainfall through a Distributed Hydrologic Model. *J. Hydrometeorol.* **2010**, *11*, 520–532. [CrossRef]
25. Li, Y.; Grimaldi, S.; Walker, J.; Pauwels, V. Application of Remote Sensing Data to Constrain Operational Rainfall-Driven Flood Forecasting: A Review. *Remote Sens.* **2016**, *8*, 456. [CrossRef]
26. Maggioni, V.; Massari, C. On the performance of satellite precipitation products in riverine flood modeling: A review. *J. Hydrol.* **2018**, *558*, 214–224. [CrossRef]
27. Mei, Y.; Anagnostou, E.N.; Shen, X.; Nikolopoulos, E.I. Decomposing the satellite precipitation error propagation through the rainfall-runoff processes. *Adv. Water Resour.* **2017**, *109*, 253–266. [CrossRef]
28. Rossi, M.; Luciani, S.; Valigi, D.; Kirschbaum, D.; Brunetti, M.; Peruccacci, S.; Guzzetti, F. Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data. *Geomorphology* **2017**, *285*, 16–27. [CrossRef]
29. Marra, F.; Destro, E.; Nikolopoulos, E.I.; Zoccatelli, D.; Creutin, J.D.; Guzzetti, F.; Borga, M. Impact of rainfall spatial aggregation on the identification of debris flow occurrence thresholds. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 4525–4532. [CrossRef]
30. Mei, Y.; Nikolopoulos, E.; Anagnostou, E.; Zoccatelli, D.; Borga, M. Error Analysis of Satellite Precipitation-Driven Modeling of Flood Events in Complex Alpine Terrain. *Remote Sens.* **2016**, *8*, 293. [CrossRef]
31. Kugler, Z.; De Groeve, T. *The Global Flood Detection System*; JRC Scientific and Technical Reports; European Communities: Luxembourg, 2007; pp. 1–45.
32. Revilla-Romero, B.; Wanders, N.; Burek, P.; Salamon, P.; de Roo, A. Integrating remotely sensed surface water extent into continental scale hydrology. *J. Hydrol.* **2016**, *543*, 659–670. [CrossRef]
33. Revilla-Romero, B.; Thielen, J.; Salamon, P.; Groeve, T.D.; Brakenridge, G.R. Evaluation of the satellite-based Global Flood Detection System for measuring river discharge: influence of local factors. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4467–4484. [CrossRef]

34. Speirs, P.; Gabella, M.; Berne, A. A Comparison between the GPM Dual-Frequency Precipitation Radar and Ground-Based Radar Precipitation Rate Estimates in the Swiss Alps and Plateau. *J. Hydrometeorol.* **2017**, *18*, 1247–1269. [CrossRef]
35. Zhang, A.; Chen, S.; Fan, S. Comparison of Extreme Precipitation Estimation From GPM Dual-Frequency Radar and Ground-Based Radar Network in Southern China. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 4534–4537.
36. Gabella, M.; Speirs, P.; Hamann, U.; Germann, U.; Berne, A. Measurement of Precipitation in the Alps Using Dual-Polarization C-Band Ground-Based Radars, the GPM Spaceborne Ku-Band Radar, and Rain Gauges. *Remote Sens.* **2017**, *9*, 1147. [CrossRef]
37. Cannon, F.; Ralph, F.M.; Wilson, A.M.; Lettenmaier, D.P. GPM Satellite and Radar Measurements and of Precipitation and Freezing Level and in Atmospheric Rivers: Comparison with Ground-Based Radars and Reanalyses. *J. Geophys. Res. Atmos. Forest.* **2017**, *122*, 12747-12764. [CrossRef]
38. Gao, J.; Tang, G.; Hong, Y. Similarities and Improvements of GPM Dual-Frequency Precipitation Radar (DPR) upon TRMM Precipitation Radar (PR) in Global Precipitation Rate Estimation, Type Classification and Vertical Profiling. *Remote Sens.* **2017**, *9*, 1142. [CrossRef]
39. Biswas, S.K.; Le, M.; Chandrasekar, V. Identification of Snow from GPM-DPR observations and cross validation with S-Band Ground Radar dual polarization measurements. In Proceedings of the 32 nd URSI GASS, Montreal, QC, Canada, 19–26 August 2017.
40. Skofronick-Jackson, G.; Hudak, D.; Petersen, W.; Nesbitt, S.W.; Chandrasekar, V.; Durden, S.; Gleicher, K.J.; Huang, G.J.; Joe, P.; Kollias, P.; et al. Global Precipitation Measurement Cold Season Precipitation Experiment (GCPEX): For Measurement's Sake, Let It Snow. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 1719–1741. [CrossRef]
41. You, Y.; Wang, N.Y.; Ferraro, R.; Rudlosky, S. Quantifying the Snowfall Detection Performance of the GPM Microwave Imager Channels over Land. *J. Hydrometeorol.* **2017**, *18*, 729–751. [CrossRef]
42. Casella, D.; Panegrossi, G.; Sanò, P.; Marra, A.C.; Dietrich, S.; Johnson, B.T.; Kulie, M.S. Evaluation of the GPM-DPR snowfall detection capability: Comparison with CloudSat-CPR. *Atmos. Res.* **2017**, *197*, 64–75. [CrossRef]
43. Wen, Y.; Behrangi, A.; Lambrigtsen, B.; Kirstetter, P.E. Evaluation and Uncertainty Estimation of the Latest Radar and Satellite Snowfall Products Using SNOTEL Measurements over Mountainous Regions in Western United States. *Remote Sens.* **2016**, *8*, 904. [CrossRef]
44. 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]
45. Rysman, J.F.; Panegrossi, G.; Sanò, P.; Marra, A.; Dietrich, S.; Milani, L.; Kulie, M. SLALOM: An All-Surface Snow Water Path Retrieval Algorithm for the GPM Microwave Imager. *Remote Sens.* **2018**, *10*, 1278. [CrossRef]
46. Gampe, D.; Ludwig, R. Evaluation of Gridded Precipitation Data Products for Hydrological Applications in Complex Topography. *Hydrology* **2017**, *4*, 53. [CrossRef]
47. *RADOLAN/RADVOR Hoch aufgelöste Niederschlagsanalyse und –Vorhersage auf der Basis Quantitativer Radar und Ombrometerdaten für and grenzüberschreitende Fluss-Einzugsgebiete von Deutschland im Echtzeitbetrieb Beschreibung des Kompositformats Version 2.4.3*; Technical Report; Deutscher Wetterdienst, Abteilung Hydrometeorologie: Offenbach, Germany, 2018.
48. Bartels, H. *Projekt RADOLAN. Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von Automatischen Bodenniederschlagsstationen (Ombrometer)*; Technical Report; Deutscher Wetterdienst, Hydrometeorologie: Offenbach, Germany, 2004.
49. Meyer, H.; Kühnlein, M.; Appelhans, T.; Nauss, T. Comparison of Four Machine Learning Algorithms for Their Applicability in Satellite-Based Optical Rainfall Retrievals. *Atmos. Res.* **2016**, *169*, 424–433. [CrossRef]
50. Kühnlein, M.; Appelhans, T.; Thies, B.; Nauss, T. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—A random forests-based approach applied to MSG SEVIRI. *Remote Sens. Environ.* **2014**, *141*, 129–143. [CrossRef]
51. Bronstert, A.; Agarwal, A.; Boessenkool, B.; Crisologo, I.; Fischer, M.; Heistermann, M.; Köhn-Reich, L.; López-Tarazón, J.A.; Moran, T.; Ozturk, U.; et al. Forensic hydro-meteorological analysis of an extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany. *Sci. Total Environ.* **2018**, *630*, 977–991. [CrossRef]

52. Doycheva, K.; Horn, G.; Koch, C.; Schumann, A.; König, M. Assessment and weighting of meteorological ensemble forecast members based on supervised machine learning with application to runoff simulations and flood warning. *Adv. Eng. Inform.* **2017**, *33*, 427–439. [CrossRef]
53. Fischer, F.; Hauck, J.; Brandhuber, R.; Weigl, E.; Maier, H.; Auerswald, K. Spatio-temporal variability of erosivity estimated from highly resolved and adjusted radar rain data (RADOLAN). *Agric. For. Meteorol.* **2016**, *223*, 72–80. [CrossRef]
54. Winterrath, T.; Brendel, C.; Hafer, M.; Junghänel, T.; Klameth, A.; Walawender, E.; und Andreas Becker, E.W. *Erstellung Einer Radargestützten Niederschlagsklimatologie*; Berichte des Deutschen Wetterdienstes; Deutschen Wetterdienstes: Offenbach am Main, Germany, 2017; Volume 251, pp. 1–71.
55. Winterrath, T.; Rosenow, W.; Weigl, E. On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management. *Weather Radar Hydrol.* **2012**, *351*, 323–329.
56. Keupp, L.; Winterrath, T.; Hollmann, R. *Use of Weather Radar Data for Climate Data Records in WMO Regions IV and VI*; Technical Report, WMO CCl TT-URSDCM; WMO: Geneva, Switzerland, 2017.
57. Richter, D. *Ergebnisse methodischer Untersuchungen zur Korrektur des Systematischen Meßfehlers des Hellmann-Niederschlagmessers*; Deutschen Wetterdienstes: Offenbach am Main, Germany, 1995; Volume 194, pp. 1–93.
58. World Meteorological Organization. *Guide to Meteorological Instruments and Methods of Observation, WMO-No. 8*; World Meteorological Organization: Geneva, Switzerland, 2017.
59. Becker, A.; Finger, P.; Meyer-Christoffer, A.; Rudolf, B.; Schamm, K.; Schneider, U.; Ziese, M. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth Syst. Sci. Data* **2013**, *5*, 71–99. [CrossRef]
60. Ullrich, P.A.; Taylor, M.A. Arbitrary-Order Conservative and Consistent Remapping and a Theory of Linear Maps: Part I. *Mon. Weather Rev.* **2015**, *143*, 2419–2440. [CrossRef]
61. Chen, C.T.; Knutson, T. On the Verification and Comparison of Extreme Rainfall Indices from Climate Models. *J. Clim.* **2008**, *21*, 1605–1621. [CrossRef]
62. Accadia, C.; Mariani, S.; Casaioli, M.; Lavagnini, A.; Speranza, A. Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average method on high-resolution verification grids. *Weather Forecast.* **2003**, *18*, 918–932. [CrossRef]
63. Chen, C.J.; Senarath, S.U.S.; Dima-West, I.M.; Marcella, M.P. Evaluation and restructuring of gridded precipitation data over the Greater Mekong Subregion. *Int. J. Climatol.* **2016**, *37*, 180–196. [CrossRef]
64. Jones, P.W. First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon. Weather Rev.* **1999**, *127*, 2204–2210. [CrossRef]
65. Jones, P.W. *A User's Guide for SCRIP: A Spherical Coordinate Remapping and Interpolation Package, Version 1.5*; Los Alamos National Laboratory: Los Alamos, NM, USA, 1998.
66. Schulzweida, U. *CDO User Guide*; Technical Report; MPI for Meteorology: Hamburg, Germany, 2018.
67. Kalognomou, E.A.; Lennard, C.; Shongwe, M.; Pinto, I.; Favre, A.; Kent, M.; Hewitson, B.; Dosio, A.; Nikulin, G.; Panitz, H.J.; et al. A Diagnostic Evaluation of Precipitation in CORDEX Models over Southern Africa. *J. Clim.* **2013**, *26*, 9477–9506. [CrossRef]
68. Diaconescu, E.P.; Gachon, P.; Laprise, R. On the Remapping Procedure of Daily Precipitation Statistics and Indices Used in Regional Climate Model Evaluation. *J. Hydrometeorol.* **2015**, *16*, 2301–2310. [CrossRef]
69. Wang, J.; Zeng, N.; Wang, M.; Jiang, F.; Chen, J.; Friedlingstein, P.; Jain, A.K.; Jiang, Z.; Ju, W.; Lienert, S.; et al. Contrasting interannual atmospheric CO_2 variabilities and their terrestrial mechanisms for two types of El Niños. *Atmos. Chem. Phys.* **2018**, *18*, 10333–10345. [CrossRef]
70. Woodcock, F. The evaluation of yes/no forecasts for scientific and administrative purposes. *Mon. Weather Rev.* **1976**, *104*, 1209–1214. [CrossRef]
71. Doswell, C.A.; Davies-Jones, R.; Keller, D.L. On summary measures of skill in rare event forecasting based on contingency tables. *Weather Forecast.* **1990**, *5*, 576–585. [CrossRef]
72. Schaefer, J.T. The critical success index as an indicator of warning skill. *Weather Forecast.* **1990**, *5*, 570–575. [CrossRef]

73. Kochendorfer, J.; Rasmussen, R.; Wolff, M.; Baker, B.; Hall, M.E.; Meyers, T.; Landolt, S.; Jachcik, A.; Isaksen, K.; Brækkan, R.; et al. The quantification and correction of wind-induced precipitation measurement errors. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 1973–1989. [CrossRef]
74. Boudala, F.S.; Isaac, G.A.; Filman, P.; Crawford, R.; Hudak, D.; Anderson, M. Performance of Emerging Technologies for Measuring Solid and Liquid Precipitation in Cold Climate as Compared to the Traditional Manual Gauges. *J. Atmos. Ocean. Technol.* **2017**, *34*, 167–185. [CrossRef]

remote sensing

MDPI

Article

Evaluation of Three-Hourly TMPA Rainfall Products Using Telemetric Rain Gauge Observations at Lai Nullah Basin in Islamabad, Pakistan

Asid Ur Rehman [1,2], Farrukh Chishtie [3,*], Waqas A. Qazi [1], Sajid Ghuffar [1], Imran Shahid [1] and Khunsa Fatima [4]

1 Department of Space Science, Institute of Space Technology (IST), Islamabad 44000, Pakistan; asid.rehman@gmail.com (A.U.R.); waqas.qazi@grel.ist.edu.pk (W.A.Q.); sajid.ghuffar@grel.ist.edu.pk (S.G.); imran.shahid@grel.ist.edu.pk (I.S.)
2 Hagler Bailly Pakistan, Islamabad 44000, Pakistan
3 SERVIR-Mekong, Asian Disaster Preparedness Center (ADPC), Bangkok 10400, Thailand
4 Institute of Geographical Information System (IGIS), National University of Science and Technology (NUST), Islamabad 44000, Pakistan; khunsafatima@gmail.com
* Correspondence: farrukh.chishtie@adpc.net; Tel.: +66-0909-29-1450

Received: 14 October 2018; Accepted: 11 December 2018; Published: 14 December 2018

Abstract: Flash floods which occur due to heavy rainfall in hilly and semi-hilly areas may prove deleterious when they hit urban centers. The prediction of such localized and heterogeneous phenomena is a challenge due to a scarcity of in-situ rainfall. A possible solution is the utilization of satellite-based precipitation products. The current study evaluates the efficacy of Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) three-hourly products, i.e., 3B42 near-real-time (3B42RT) and 3B42 research version (3B42V7) at a sub-daily time scale. Various categorical indices have been used to assess the capability of products in the detection of rain/no-rain. Hourly rain rates are assessed by employing the most commonly used statistical measures, such as correlation coefficients (CC), mean bias error (MBE), mean absolute error (MAE), and root-mean-square error (RMSE). Further, a diurnal analysis is performed to authenticate TMPA's performance in specific hours of the day. In general, the results show the good capability of both TMPA products in the detection of rain/no-rain events in all seasons except winter. Specifically, 3B42V7 performed better than 3B42RT. Moreover, both products detect a high number of rainy days falsely in light rain ranges. Regarding rainfall measurements, TMPA products exhibit an overall underestimation. Seasonally, 3B42V7 underestimates rainfall in monsoon and post-monsoon, and overestimates in winter and pre-monsoon. 3B42RT, on the other hand, underestimates rainfall in all seasons. A greater MBE and RMSE are found with both TMPA rain measurements in monsoon and post-monsoon seasons. Overall, a weak correlation and high MBE between the TMPA (3B42RT, 3B42V7) and reference gauge hourly rain rates are found at a three-hourly time scale (CC = 0.41, 0.38, MBE = −0.92, −0.70). The correlation is significant at decadal (CC = 0.79, 0.77) and monthly (CC = 0.91, 0.90) timescales. Furthermore, diurnal rainfall analysis indicates low credibility of 3B42RT to detect flash flooding. Within the parameters of this study, we conclude that the TMPA products are not the best choice at a three-hourly time scale in hilly/semi-hilly areas of Pakistan. However, both products can be used at daily, yet more reliably above daily, time scales, with 3B42V7 preferable due to its consistency.

Keywords: TRMM; TMPA; 3B42; validation; rainfall; telemetric rain gauge; Lai Nullah; Pakistan

1. Introduction

Precipitation is a key phenomenon which has a direct impact on global climate patterns. Anomalies in precipitation may cause natural hazards, like droughts, floods, and landslides [1]. A natural event that occurs in a populated area is highly dangerous as it may cause loss of human lives and damage to property [2]. Regarding the climate change vulnerability index, Pakistan has been ranked 13th globally [3]. The county is prone to multiple natural hazards, like floods, droughts, earthquakes, and cyclones [3–5]. The frequency and intensity of extreme weather events have increased in the past two decades; therefore, an estimated 40% of the population is vulnerable to such disasters [3]. Rainfall, a major contributing factor at local-scale climate patterns, not only impacts weather systems, but also influences the water flow in rivers and water level in dams [3]. Monsoon rain is an expected and dominant phenomenon in Pakistan's climate [6]. The monsoon period typically ranges from July to mid-September. Most parts of the country receive rainfall in the monsoon season [7]. Due to heavy rainfall in the monsoon season, flooding in the low-lying areas along rivers is common, while hilly and semi-hilly areas experience flash floods due to steep elevation gradients [7,8]. Lai Nullah is a natural stream fed by rainwater. It passes through the twin cities, i.e., Islamabad and Rawalpindi, and inundates the surrounding areas almost every year due to heavy rainfall in the monsoon [3,6]. Lai Nullah basin receives rainfall of about 600 mm during the monsoon, which is ~60% of annual rainfall. From 1944 to 2014 (70 years), Lai Nullah basin faced more than 23 flood events [9]. Extreme flood events were recorded in 1981, 1988, 1997, and 2001. The flood in 2001 was the most catastrophic, causing 74 deaths, out of which 64 fatalities occurred in Rawalpindi [3,6,10,11]. From 1994 to 1997, Lai Nullah flooded five times, causing the loss of 84 human lives and destroying around 1000 houses [6,10,12,13]. After the 2001 floods, intermittent moderate level floods have been reported. A flash flood was reported by Hayat (2007) [4] in Lai Nullah due to moderate intensity rain for 6 h on 20th March 2007. Afzal et al. [14] reported a severe rainfall spell on 5th July 2008 when 104 mm rainfall was recorded in 100 min, and an overall 162 mm in 5 h. Consequently, densely populated low-lying areas near Lai Nullah faced a flood-like situation, causing three causalities and wrecking a massive amount of property. In 2014, a flood occurred when ~280 mm rainfall was recorded in a day, which severely inundated low-lying areas [6]. Each year, after the flooding, people living along Lai Nullah face numerous crises, which include loss of life and property [15]. There is a dire need for accurate and high spatio-temporal precipitation data to minimize risks to people's lives and properties due to regular flash floods in Lai Nullah. Such measurements are not only important to predict rapidly developing hydro-meteorological phenomena like a flash flood, but also significant for sustainable planning and the management of water resources, agriculture lands, urban areas, and ecological systems [16,17].

Conventionally, rainfall gauges are used to observe rainfall. Presently, these are the only instruments available to measure rainfall accumulation directly. Over the last few decades, the number of rainfall gauges has increased tremendously, but still, their density does not meet scientific needs. Furthermore, rainfall is a heterogeneous phenomenon which varies at different spatio-temporal scales. Rain gauges only record local-scale information, which may lead to potential errors when interpolation is applied at larger scales, especially in hilly areas [18]. In addition, a spatial distribution of rain gauges is sparse in remote, undeveloped areas and regions with complex terrain [19]. An alternative approach to estimating precipitation is a weather radar, which can reveal the precipitation patterns and storm structure along with real-time monitoring at a very high spatial and temporal resolution. Limited coverage of weather radars and the associated complex data processing that is also prone to errors is another challenge [1]. Thus, recording precise precipitation at a fine spatial and temporal resolution, from a regional to global scale, is a contemporary issue for researchers [1,16,20].

Advances in remote sensing technologies have capacitated the scientific community to employ satellite-based precipitation products at varied spatial and temporal resolutions [16]. Satellites such as the Tropical Rainfall Measuring Mission (TRMM) [21], Climate Prediction Center MORPHing technique (CMORPH) [22], and Precipitation Estimation from Remotely Sensed Information Using

Artificial Neural Networks (PERSIANN) [23] are delivering free data products to the community. These products are indeed valuable for hydrological modeling and predictions, flood forecasting, land surface derivation models, reservoir operations and risk analysis, resource management and monitoring, drought early warning, and the validation of numerical models for regions with a paucity of weather stations. Typically, these rainfall products have a high spatial and temporal resolution, and good spatial and temporal coverage. [1,24,25]

TRMM products are widely used among researchers and hydrologists, and have been applied globally in various precipitation assessment, and hydrological and flood inundation modeling studies [25–39]. TRMM was the first space mission with a special focus on measuring tropical and subtropical rainfall through visible/infrared, and active and passive microwave sensors [40]. Primarily, the three rainfall sensors mounted on TRMM were the Visible and Infrared Scanner (VIRS), Precipitation Radar (PR), and TRMM Microwave Imager (TMI) [16,41].

While the advantages of such satellite-based products are multifarious, precipitation measurements that are done through indirect methods associated with remote sensing techniques can potentially induce large uncertainties and errors. These uncertainties associated with the precipitation retrieval algorithm can generate errors in rainfall estimation, which can further propagate into hydrological modeling. Therefore, it is essential to evaluate the rainfall retrieval capability of a satellite sensor before its application. Such an assessment can guide users about applications of products and impacts of accompanied errors [1,17,42].

Many studies have been undertaken globally to validate Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) at a daily time scale. For example, Wang et al. [1] evaluated TMPA 3B42 version 7 and Integrated Multi-satellite Retrievals for GPM (IMERG) at different time scales and found a very high correlation of TMPA 3B42 version 7 with rain gauge observations at a yearly time scale, but a weak linear relationship (correlation coefficient <0.5) at a daily time scale. A couple of studies [43,44] evaluated TMPA 3B42 version 6 in Saudi Arabia and Nepal. The studies showed an overall underestimation in Nepal and a typical overestimation in Saudi Arabia. Moreover, the performance of TMPA 3B42 version 6 in mountainous regions was poor compared to plain areas. A similar study [45] conducted in Bangladesh showed that TMPA 3B42 version 5 overestimated rainfall in most of the areas in pre-monsoon and post-monsoon seasons and underestimated in the monsoon season, especially in heavily wet areas. Overall, TMPA 3B42 version 5 underestimated rainfall in wet areas and overestimated rainfall in dry areas. Cai et al. [46] also found a weak correlation ($r = 0.5$) between TMPA 3B42 version 7 and rain gauge data, but the correlation was high at monthly and annual time scales. Another study [47] used different temporal as well as spatial scales to evaluate TMPA 3B42 and found a modest correlation at a daily scale, but a good relationship ($r > 0.8$) at monthly timescale. Additionally, the study revealed no significant improvement in the performance of TMPA when evaluated at a 1×1 degree grid cell by merging 0.25×0.25 degree grid cells.

Though TMPA 3B42 products have been extensively assessed at daily and longer time scales, their evaluation at a sub-daily time scale is quite rare. A recent study [17] compared TMPA 3B42 research version with rain gauge data at Khorasan Razavi province in northeastern Iran and observed a correlation coefficient value of 0.36 for average three-hourly rainfall. The study also discovered an underestimation of TMPA 3B42 research version at a 3-h timescale. Another study [48] conducted in Greece evaluated TMPA 3B42 with rain gauge data at a high temporal resolution of 3 h and found a correlation coefficient value slightly less than 0.3 between both sources. Nevertheless, this correlation increased with aggregated time intervals, i.e., 6 h, 12 h, 24 h, and 48 h, and gained a peak value of 0.6 at a maximum aggregated interval. Moreover, the study found a decrease in mean absolute error and root mean square error when moving from a low aggregated interval to a high aggregated interval. Yuan et al. [49] investigated the diurnal characteristics of TRMM 3B42 in central eastern China by using rain gauge data and noticed some dispersed patterns. TRMM 3B42 version 6 hourly intensities were also appraised by Scheel et al. [47] for a very short period of three months in the Central Andes region. Results regarding rain detection capabilities were tantamount to daily results, but the correlation was

significantly lower (0.01) compared to the correlation (r = 0.55) with ground observations at a daily timescale. In another study [50] conducted over the Korean Peninsula, the correlation coefficient value was found to be 0.47 when 3B42 V6 was evaluated against automatic weather station data, whereas the 3B42 product slightly overestimated rain rates. Villarini [51] compared two pixels of the three-hourly rainfall product of TMPA 3B42 with rain gauge observations over the metropolitan area and observed a mean correlation coefficient value of 0.38 between both sources. In another study [52] on a regional scale, a mean correlation coefficient value of 0.45 between TMPA 3B42 and ground rainfall observations was noticed for cold and warm seasons.

Many studies have been carried out for Pakistan regarding the validation of TMPA 3B42 products. Nonetheless, these studies only validated TMPA 3B42 products at daily or above daily time scales. In a study [16], surface rainfall gauge and Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) data were used to evaluate TMPA daily rain. The study divulged an overestimation of TMPA 3B42 version 7 in pre-monsoon and monsoon seasons, while an underestimation in post-monsoon and winter seasons. Likewise, a good agreement between TMPA 3B42 version 7 and ground observations was observed in plain and medium elevation areas. Hussain et al. [53] concluded in their study that the performance of TMPA 3B42 was better in mountainous regions than the arid region. Also, TMPA 3B42 overestimated rainfall in glacier regions. Anjum et al. [54] assessed TMPA 3B42 real-time and post-real-time products for the inclement flood event of 2010 in the mountainous region. The TMPA post-real-time product performed well in comparison to the real-time product and generally captured the spatial patterns of precipitations; albeit both products were unable to capture the temporal variation of the storm. TMPA 3B42 v6 and v7 were compared by Anjum et al. [55] for the Swat watershed. Overall, TMPA 3B42 version 7 performed better. The algorithms performed well at the monthly scale, but poorly at the daily and seasonal scale. Another study [56] revealed that TMPA 3B42 underestimated rainfall in intense monsoon regions, but overestimated in southern parts of the Indus basin. Three different satellite algorithms, including the TMPA 3B42 real-time product, were analyzed by Sadiq et al. [57]. The TMPA performed well in high intensity rainfall events, but highly overestimated for light rain rates. This overestimation was also observed in mountains regions in north and coastal areas in the south of the Indus basin.

Our present study comprises the validation and performance assessment of both TMPA 3-hourly rainfall products, i.e., 3B42 near-real-time (3B42RT hereafter) and 3B42 research version (3B42V7 hereafter) hourly rain rates at various time scales for the period 2007 to 2010 in Lai Nullah basin in Islamabad, Pakistan. The major objective of the study is to evaluate TMPA's rainfall rates (mm/h) at 3-h time intervals by using 10-min temporal resolution telemetric rain gauge (TRG) observations. Validation of TMPA products at the finest time scale, i.e., 3-h, is inadequate around the globe. One possible reason for this is that the in-situ rainfall data is not sufficiently available at the sub-daily timescale. Similarly, in Pakistan, the validation of TMPA products at the sub-daily scale has not been executed yet, and the current study is the first of its kind to evaluate TMPA products at a 3-h timescale. The motivation behind this validation study is to examine the feasibility of 3B42RT's three-hourly rain rates for the prediction of heterogeneous and localized hydro-meteorological phenomenon like urban flash flooding and to determine the viability of 3B42V7's sub-daily rain rates for water-related modeling. Furthermore, diurnal analysis has also been conducted to verify TMPA products' behavior at specific hours of the day. The TMPA's rain rates have also been inspected at daily, decadal, and monthly timescales. Aggregated rainfall patterns of both TMPA products are also analyzed at seasonal and annual scales. Such a fine-scale validation study will assist users and scientists to probe the operability of the TRMM 3B42 3-h products before their potential usage at a high temporal resolution (3-h) in the localized context of Pakistan.

2. Study Area, Datasets, and Methodology

2.1. Study Area

The study area is Lai Nullah basin, which covers a major portion of Islamabad and Rawalpindi, also known as the 'Twin cities', in the northern part of Pakistan. This basin is located at 33°45′36″ to 33°33′01″ North and 72°55′16″ to 73°6′27″ East (Figure 1).

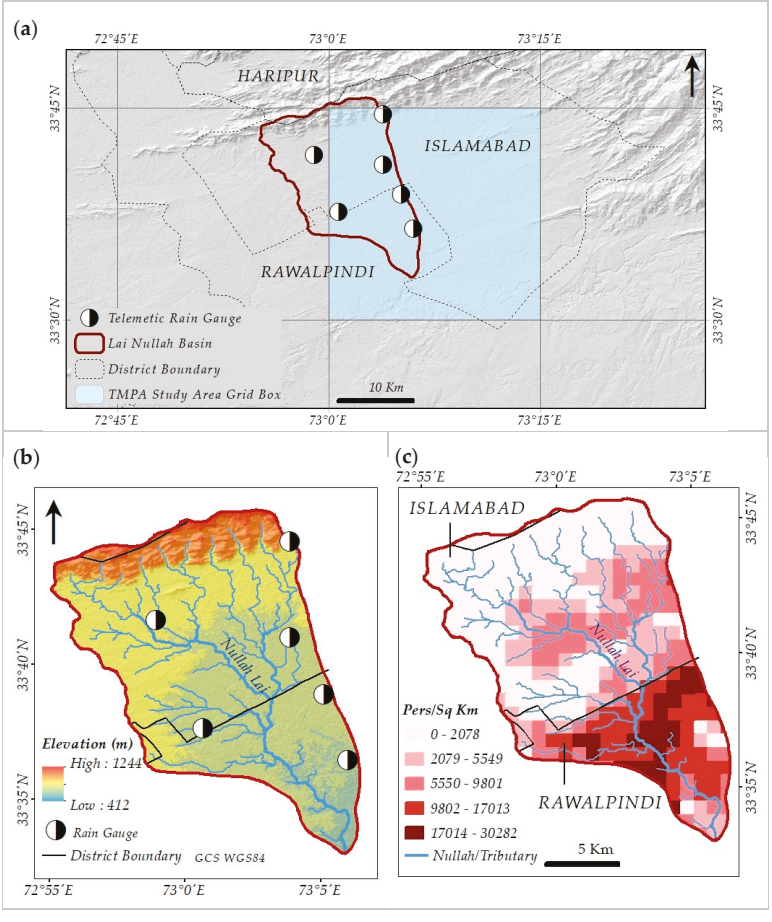


Figure 1. Map of study area showing (**a**) location of TRGs and TMPA pixel area, (**b**) elevation, and (**c**) population density (persons per square kilometer). (Data sources: ASTER 30-m Digital Elevation Model for topography, LandScan 2011 for population density).

The total area of the basin is approximately 230 km^2. The area of the upper portion of the basin which covers Islamabad is approximately 159.6 km^2, i.e., ~69% of the total catchment area of Lai Nullah basin. The lower part of the basin covers almost 70.5 km^2 of Rawalpindi city and its surroundings [6,10,58]. The lowest point of Lai Nullah is 420 m above mean sea level (a.m.s.l.) at the confluence of Soan River, while the highest elevation of the basin is 1200 m a.m.s.l. at the top of the mountain of Margalla hills.

2.2. Telemetric Rain Gauge Observations

An in-situ rainfall dataset for a period of approximately four years (1 April 2007 to 31 December 2010) at a 10 min temporal resolution was obtained from the Pakistan Meteorological Department (PMD)—the national weather forecasting agency of Pakistan. The surface rainfall observations have been recorded by telemetric rain gauges (TRGs). PMD, with the technical help of Japan International Cooperation Agency (JICA), installed these TRGs in Lai Nullah basin for a flood prediction and early warning system. These telemetric rain gauges are tipping bucket rainfall gauges which have a sensitivity of 0.1 mm [12]. Tipping bucket (TB) rain gauges may include systematic as well as random errors. Systematic error may include calibration error, wind-induced error, and sampling error. PMD follows the World Meteorological Organization (WMO) guidelines [59] on quality control for automated weather stations, which include telemetric rain gauges. Figure 1a shows the location while Table 1 provides the exact coordinates and elevation of TRGs. Out of the six gauges, only two rain gauges are in Rawalpindi, while the remaining are stationed in Islamabad city.

Table 1. PMD telemetric rain gauges' specifications.

SN	Name	District	Coordinates		Elevation (m)
			Longitude (E)	Latitude (N)	
1	Saidpur	Islamabad	73°03′51″	33°44′33″	668
2	Golra	Islamabad	72°58′55″	33°41′38″	556
3	PMD	Islamabad	73°03′51″	33°40′59″	531
4	Bokra	Islamabad	73°00′39″	33°37′38″	528
5	RAMC	Rawalpindi	73°05′07″	33°38′53″	508
6	Chaklala	Rawalpindi	73°05′60″	33°36′27″	506

2.3. TMPA

The Tropical Rainfall Measuring Mission was launched on 27 November 1997, from the Tanegashima Space Center in Japan. Various weather-related instruments like a TRMM microwave imager (TMI), and visible and infrared sensor (VIRS) were mounted on the satellite [43]. The TMPA is one of several products offered by TRMM, and it is a fine scale product which has a spatial resolution of 0.25° × 0.25° at a 3-h temporal scale [1,21]. The product covers an area between 50°N–50°S. The construction of TMPA involves multi-sensor analysis techniques which utilize synergy between the passive microwave retrievals of rainfall with the geostationary infrared observations [41]. The TMPA is further categorized as real-time (3B42-RT) and post-real time (3B42 research version) [21,60]. Both versions have the same spatial and temporal coverage; the only difference is that the research version is biased adjusted by using monthly rain gauge data [21,61]. The current study includes both TMPA versions for the validation.

TMPA 3B42 products are developed in a sequence which starts from a collection of rainfall-related passive microwave data from various low earth orbit (LEO) microwave sensors, including TMI, Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E), Special Sensor Microwave Imager (SSM/I), and Advanced Microwave Sounding Unit-B (AMSU-B). Each of the passive microwave sensor's measurements are converted into precipitation and then averaged to a spatial grid of a 0.25° cell size with a temporal resolution of 3-h (±90 min from central time slot). Best estimates are calculated by using some statistical methods for coincident data. However, it should be noted that in LEO, each satellite gives a very dispersed sampling of precipitation in time and space. All other microwave sensors are calibrated by using TRMM combine instruments (TCI) and TMI. In the next step, infrared measurements are collected from the international constellation of geosynchronous earth orbit (GEO) satellites. Then, GEO's IR data, which offers an exceptional time-space coverage, is processed at a 0.25° spatial grid and 3-h temporal scale. Again, microwave TCI data is used to calibrate IR estimates. Then, microwave and infrared estimates are merged to form a single best estimate.

In first preference, 0.25° grid cells are assigned microwave-based rainfall estimates, and the remaining gaps are filled with IR-based precipitating estimates. This step provides near-real-time estimates of precipitation labeled as the TMPA 3B42 near-real time product. At the final stage, three-hourly merged rain estimates are summed over a calendar month to construct the monthly multi-satellite (MS) product. This product is then combined with monthly rain gauge data to form a post-real time satellite-gauge (SG) product, which is labeled as TRMM's research version product (3B43). In the end, the SG/MS ratios are calculated on each cell of a 0.25° grid and then applied to every three-hourly field of the month to construct a research version 3B42 product [21]. Version 7 of the TMPA 3B42 product was released in May 2012 after applying some major revisions in the TMPA 3B42 version 6. However, very soon after this, some errors were identified in TMPA 3B42V7; therefore, V7 was again revisited, and a newly revised version was released in December 2012 [62].

2.4. Methodology

Validation of precipitation data at a sub-daily scale is a big challenge as standard observations are taken at a daily scale [52]. However, for the current study, reference in-situ rainfall observations are available at a very high temporal resolution of 10 min. Therefore, such high temporal reference data offers a unique opportunity to assess TMPA's hourly rain rates at its finest temporal resolution, i.e., 3-h, and then the other time scales like daily, decadal, and monthly.

In-situ reference data is taken from the telemetric rain gauges, which are point-based observations. On the other hand, a satellite averages rainfall over a specific area or pixel. One way is to interpolate reference point data for an area-to-area comparison, but such a comparison may cause additional error due to interpolation techniques [1]. Thus, the point-to-pixel comparison method has been considered for the current study. TMPA is a grid-based product; therefore, the location of each pixel for rain gauge stations remains fixed. Spatial overlays revealed that rain gauges are quite closely distributed and five out of six rain gauges fall within a single pixel. Villarini [51] suggests using at least five rain gauges for computing areal averages at high spatial and temporal resolutions. Therefore, non-zero averaged rain rates (mm/h) of five rain gauges are calculated for the comparison with satellite-based rainfall measurements. The use of more than a distributed single station against one TMPA grid cell may minimize the well-known issue of unrepresentative error, i.e., point-based precipitation observations are likely to be unrepresentative of satellite estimated areal values [20]. But the number of rain gauges that is suitable to represent a grid cell is dependent on the rain pattern, spatial properties, seasons, and terrain. However, the correlation coefficient between satellite-based rainfall and rain gauge rainfall is increased when the number of rain gauges and accumulation intervals is increased till a certain range. Also, the error statistic is improved when more than one set of rain gauge data is used against one grid cell [63]. After careful assessment, the sixth rain gauge which is falling near the border of another pixel is discarded from the analysis. Finally, temporal collocation using a ± 5 min time window is performed to generate the satellite–gauge data pairs for comparison.

Commonly used statistical methods to assess satellite rainfall algorithms are continuous and categorical methods [20]. Continuous statistical measures related to satellite rainfall amount evaluation include mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient (CC) between estimated and observed rainfall values [20,43,44,54,64]. CC computes the linear relationship between measured and observed rainfall. MBE calculates the average difference between measured and observed rainfall. A positive MBE value indicates an overestimation, while a negative value presents an underestimation of satellite rainfall with respect to the rainfall from in-situ gauge data. MAE determines the average absolute magnitude of error, while RMSE also estimates the average error magnitude, but gives more weight to larger errors. The perfect values for these measures are 1 for the correlation coefficient, and 0 for MBE, MAE, and RMSE. The CC value is considered significant if it is equal to or greater than 0.7 [53]. Standard deviation (SD) [65] is another useful statistical measure to assess the capability of each dataset in capturing rainfall variability. A higher SD value represents more variation in rainfall from its mean

value. The performance of the TMPA algorithm has been characterized into three basic groups [43,44]: overestimation, underestimation, and approximately equal (within ±10%). Categorical statistical measures have been calculated to assess the performance of TMPA in the detection of rain/no-rain events at a given threshold. Indices like Bias, probability of detection (POB), false alarms ratio (FAR), threat score (TS), equitable threat score (ETS) and skill score (SS) [1,20,43], Refs. [44,54,66,67] have been calculated. Bias estimates the frequency of measured rainfall events to the frequency of observed rainfall events. It indicates whether a measuring system under-detects or over-detects rainy events. POD describes the ability of a rain measuring source to detect rain events correctly. FAR represents the cases when a rain measuring source detects rain when there is no rain detected. TS/CSI, ETS, and SS demonstrate the overall fraction of events which are correctly detected by a rain measuring source. The difference between ETS and TS is that the ETS is an adjusted index for hits associated with random chance. Perfect scores for Bias and FAR are 0 and 1 for the POD, TS, ETS, and SS. More details of categorical and continuous statistical measures are given in Table 2. The diurnal variations of precipitation captured by TRG and TMPA products are assessed by using the frequency of precipitation occurrence, the precipitation intensity, and the precipitation amount. Precipitation frequency is defined as the number of rainy events divided by the total number of events. Precipitation intensity is defined as the mean rates averaged over the precipitating events, and the amount of precipitation is accumulated precipitation [49,68–70].

Table 2. List of the statistical indices used in the evaluation and validation of TMPA Products.

Type	Statistical Index	Equation	Perfect Value	Description
Categorical Statistical Measures	Bias	$\text{Bias} = \frac{H+FA}{H+M} - 1$	0	H = hits (rain at both) FA = false alarms (rain at measured and no-rain at observed) M = misses (no-rain at measured and rain at observed) Z = zeros (no-rain at both) N = total events
	Probability of Detection (POD)	$\text{POD} = \frac{H}{H+M}$	1	
	False Alarm Ratio (FAR)	$\text{FAR} = \frac{FA}{H+FA}$	0	
	Threat Score (TS)	$\text{TS} = \frac{H}{H+M+FA}$	1	
	Equitable Threat Score (ETS)	$\text{ETS} = \frac{H-H_{Rand}}{H+M+FA-H_{Rand}}$ $H_{Rand} = \frac{(H+M)\times(H+FA)}{N}$		
	Skill Score (SS)	$\text{SS} = \frac{Z\times H-FA\times M}{Z+FA\times M+H}$	1	
Continuous Statistical Measures	Mean Bias Error (MBE)	$\text{MBE} = \frac{1}{n}\sum_1^n (M-O)$	0	n = number of samples, M = measured value (TMPA) O = observed value (TRG), X = data element (X = M for TMPA products and X = O for TRG) Bar on variable shows mean value
	Mean Absolute Error (MAE)	$\text{MAE} = \frac{1}{n}\sum_1^n \lvert M-O\rvert$	0	
	Root Mean Square Error (RMSE)	$\text{RMSE} = \sqrt{\frac{\sum_1^n (M-O)^2}{n}}$	0	
	Correlation Coefficient (CC)	$\text{CC} = \frac{\sum_1^n (M-\overline{M})\,(O-\overline{O})}{\sqrt{\sum_1^n (M-\overline{M})^2}\sqrt{\sum_1^n (O-\overline{O})^2}}$	1	
	Standard Deviation (SD)	$\text{SD} = \sqrt{\frac{\sum_1^n (X-\overline{X})^2}{n}}$	NA	

The rain/no-rain detection threshold reported for the telemetric rain gauges is 0.1 mm/h. However, a minimum non-zero value of '1 mm/h' has been found in TRG data series analysis. Hence, the data pairs with the rain gauge rainfall value '0' and TMPA rainfall values less than 1 mm/h are excluded from categorical analysis. Overall, the rain rate threshold of 0.1 mm/h has been used for both datasets to consider rain/no-rain events for calculating categorical, as well as continuous, statistics measurements. Moreover, data pairs where TMPA products and rain gauge have zero rainfall values are excluded during calculations of continuous statistics, i.e., CC, MBE, MAE, and RMSE.

The comparison of TMPA hourly rain rates with the rain gauge observations is carried out at a 3-h time interval. Furthermore, mean hourly rain rates are computed for daily, decadal, and monthly comparisons. TMPA hourly rain rates are also evaluated in different seasons. Seasons are defined as the winter season (November-February), pre-monsoon season (March-May), monsoon season (June-August), and post-monsoon season (September-October) [71]. The mean maximum temperature of Islamabad is 40 °C during June (monsoon) and the mean minimum temperature

is 3 °C during January (winter) [72]. The annual mean temperature of Islamabad is 29 °C, and annual mean precipitation is 1140 mm [73]. The monsoon receives a major portion of annual rainfall (~60%). The monsoon precipitation system which mainly originates from the Bay of Bengal enters Pakistan from its southern (Arabian sea), south-eastern (India), and north-eastern boundaries (Nepal) [56,57,74]. Due to heavy rainfall in the monsoon, different parts of Pakistan face floods in this season. Pre-monsoon and post-monsoon seasons are considered warm seasons, with a relatively low rainfall amount in these seasons. Islamabad often receives moderate to heavy rainfall showers in winter months due to a phenomenon called western disturbances or winter westerlies. Western disturbances originate from the Mediterranean region and enter Pakistan through Afghanistan at its north-western boundaries [56,57,75].

In addition to rain rates validation, monthly, seasonal, and annual rainfall amounts are also aggregated to compare both datasets.

3. Analysis and Results

3.1. TMPA Products Evaluation for Entire Study Period

3.1.1. TMPA Performance Assessment Using Categorical Indices

The values of performance indices (as per Table 2) at a 3-h interval have been calculated and are shown in Figure 2. The Bias score for the entire period is −0.18 for 3B42RT and −0.11 for 3B42V7. The values of probability of detection (POD) and false alarm ratio (FAR) are 0.59 and 0.28, respectively, for 3B42RT, and 0.65 and 0.27, respectively, for 3BV7. The threat score (TS) or critical success index (CSI) measures the fraction of all rain events that were correctly identified and its ideal score should be '1'. In the current study, the calculated threat score is 0.48 and 0.52 for 3B42RT and 3B42V7, respectively. The equitable threat score (ETS), which estimates the skill of a measuring source relative to chance, has low values compared to TS. The skill score (SS) is 0.58 for 3B42RT and 0.64 for 3B42V7.

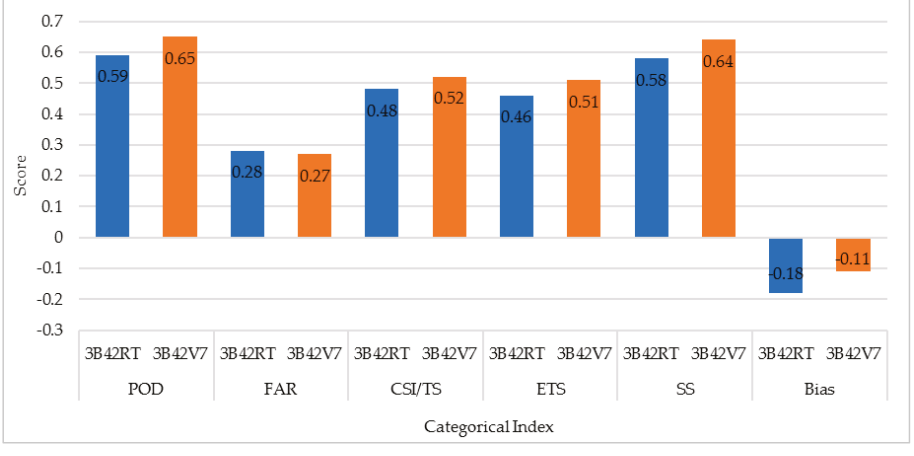


Figure 2. TMPA categorical skill scores at the three-hourly time scale.

3.1.2. TMPA Performance as a Function of Rainfall Intensity

Skill scores for different performance indices have been calculated at various rainfall intensities to check TMPA's performance in the detection of events at different rain intensities. The rain threshold to detect a rain/no-rain event is 0.1 mm/h for all rain intensity brackets. As shown in Figure 3a,b, for light rain events, both TMPA products have a high positive bias score (>0.9) and high FAR values,

which indicate that the TMPA products erroneously detect a high number of rainy days. Overall, TMPA products demonstrate a better performance for high-intensity rain events. It should also be noted that the higher the rainfall intensities, the smaller the number of data samples available for skill scores calculation. Figure 3c displays a TRG, 3B42RT, and 3B42V7 histogram which shows rainfall distribution frequencies. Histograms of all three datasets are right-skewed, and around 75% of the events fall under 0–2 mm/h rainfall intensities.

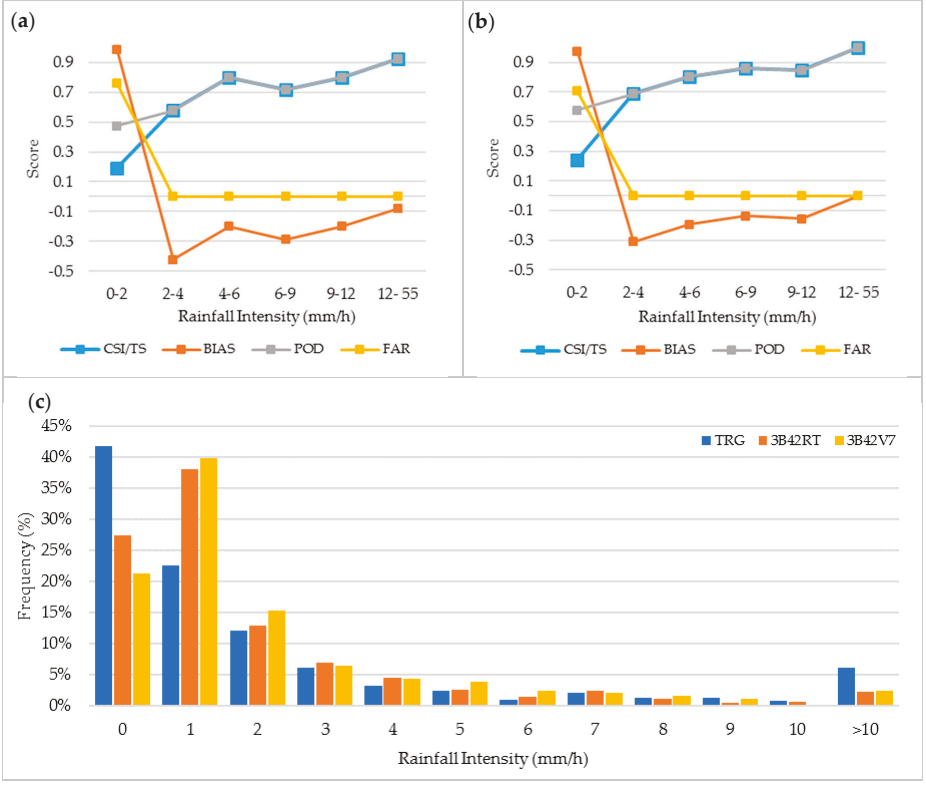


Figure 3. Performance indices as a function of an increasing rain intensity for (**a**) 3B42RT and (**b**) 3B42V7. (**c**) The number of event occurrences in percentage at various rainfall thresholds.

TMPA's performance in the detection of various rainfall intensities has been examined by classifying TRG rain rates into different rain categories (Table 3). TRG rain intensities have been categorized as light rain, moderate rain, heavy rain, very heavy rain, and extreme rain [76]. Figure 4a,b describes how 3B42RT and 3B42V7 classify rain events with reference to TRG classification. 3B42RT detects 58.5% 'light rain' events as 'no rain' and only 21.6% 'light rain' events are detected correctly, while 19.9% of 'light rain' events are detected as 'moderate' and 'heavy rain' events. Similarly, 3B42RT detects 25.9% 'moderate rain' events, 21.7% 'heavy rain events', and only 9.5% 'very heavy rain' events correctly. 3B42RT detects all 'extreme events' as 'heavy rain' intensity rain events. 3B42V7, on the other hand, detects 51.1% 'light rain' events as 'no rain' and only 25,6% 'light rain' events are detected correctly, while 23.3% of 'light rain' events are detected as 'moderate' and 'heavy rain' events. Similarly, 3B42V7 detects 31.3% 'moderate rain' events, 33.7% 'heavy rain' events, and only 4.8% 'very

heavy' rain events correctly. TMPA detects all extreme events as very heavy rain intensity rain events (see Figure 4c).

Table 3. Classification of rainfall intensities.

SN	Rain Category	Threshold (mm/h)
1	Light rain	<1
2	Moderate rain	1–4
3	Heavy rain	4–16
4	Very heavy rain	16–50
5	Extreme rain	>50

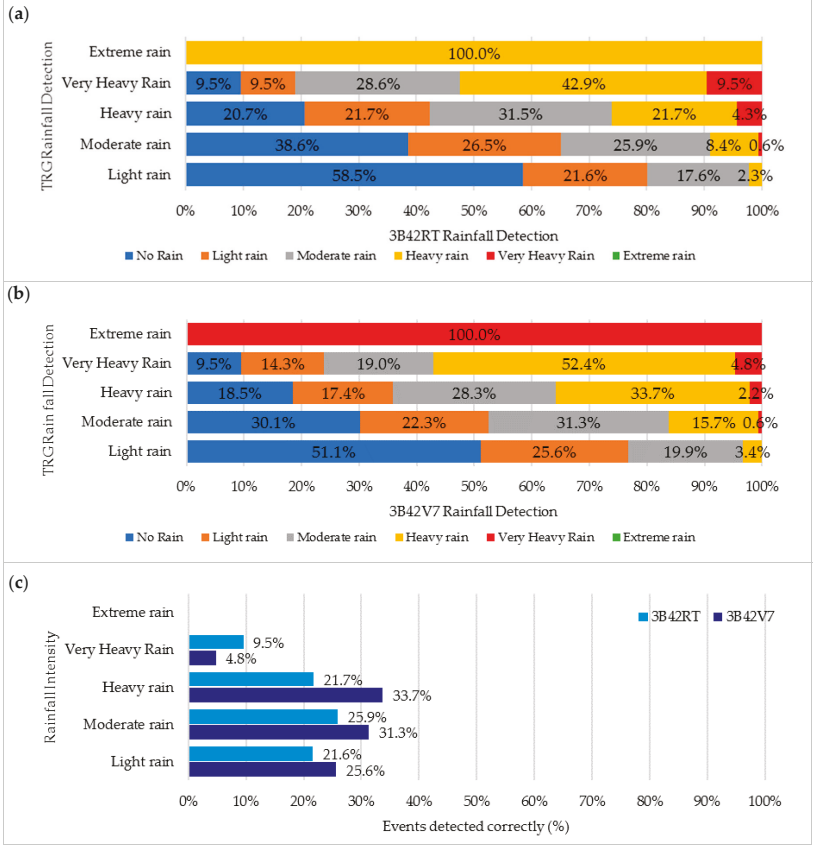


Figure 4. Performance in the detection of rainfall intensities by (**a**) 3B42RT and (**b**) 3B42V7. (**c**) Performance comparison of 3B42RT and 3B42V7.

3.1.3. TMPA Rainfall Validation

On average, 3B42RT and 3B42V7 give rain rates of 1.50 and 1.72 mm/h, respectively, at a three-hourly time scale, while the TRG rain rate is 2.42 mm/h (see Table 4). The mean bias error for 3B42RT is −0.92 mm/h, while it is −0.70 mm/h for 3B42V7. The negative sign shows an overall underestimation of rainfall by TMPA products. Mean absolute errors at a three-hourly scale are 2.47 mm/h and 2.46 mm/h, respectively. The root means square error, which gives more weight to

larger errors, has a value of 4.88 mm/h for 3B42RT and 5.01 mm/h for 3B42V7. Figure 5a,b shows the scatter plots of TRG and TMPA product measurements. The value of the correlation coefficient at a three-hourly scale is 0.41 for 3B42RT and 0.38 for 3B43V7.

Table 4. Continuous statistics for hourly rain rates (mm/h). Rain rate values at 3-h interval are instantaneous. Daily and above daily rain rates are averaged over time.

Timescale	Rain Rate (mm/h)			SD (mm/h)			MBE (mm/h)		MAE (mm/h)		RMSE (mm/h)		CC	
	TRG	3B42		TRG	3B42		3B42		3B42		3B42		3B42	
		RT	V7		RT	V7	RT	V7	RT	V7	RT	V7	RT	V7
3-h Interval	2.42	1.50	1.72	5.18	2.86	3.30	−0.92	−0.70	2.47	2.46	4.88	5.01	0.41	0.38
Mean Daily	0.40	0.30	0.35	0.65	0.51	0.60	−0.10	−0.05	0.35	0.34	0.61	0.61	0.48	0.53
Mean Decadal	0.16	0.12	0.14	0.22	0.14	0.16	−0.04	−0.02	0.08	0.07	0.14	0.14	0.79	0.77
Mean Monthly	0.14	0.11	0.12	0.17	0.11	0.12	−0.03	−0.02	0.06	0.08	0.09	0.08	0.91	0.90

Scatterplots of TMPA products vs. TRG at various time scales are shown in Figures 5 and 6. These scatterplots show that the value of the correlation coefficient between TMPA products and TRG measurements is improved when hourly rain rates are averaged at daily, decadal, and monthly time intervals. CC is the highest at the monthly time interval for both products when compared to shorter time intervals. Validation statistics, summarized in Table 4, indicate error statistics being ameliorated when shifted from brief to prolonged time intervals. Errors, i.e., MBE, MAE, and RMSE, are minimum at the monthly time interval. Standard deviation values calculated at various timescales show a quite low skill of TMPA products in capturing rainfall variability, especially at a 3-h timescale. This gap in rainfall variability is reduced at higher timescales.

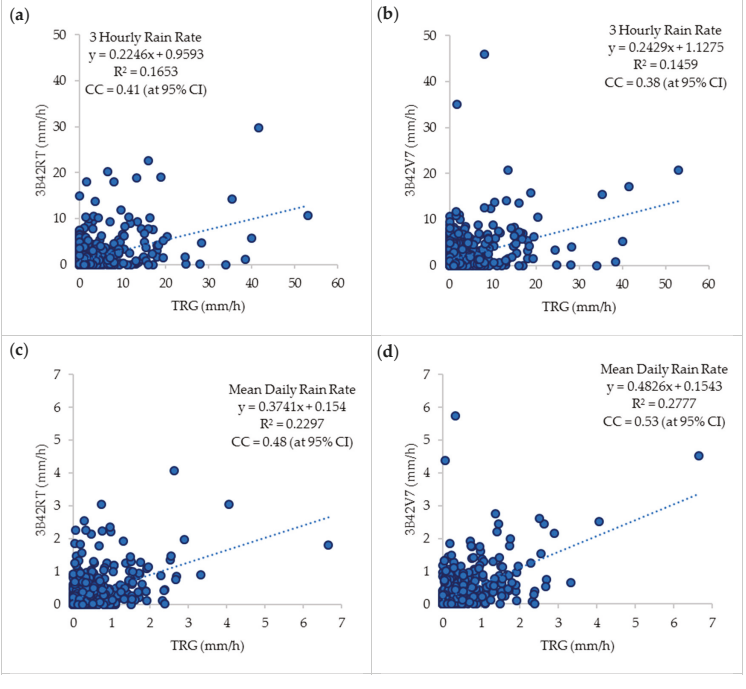


Figure 5. *Cont.*

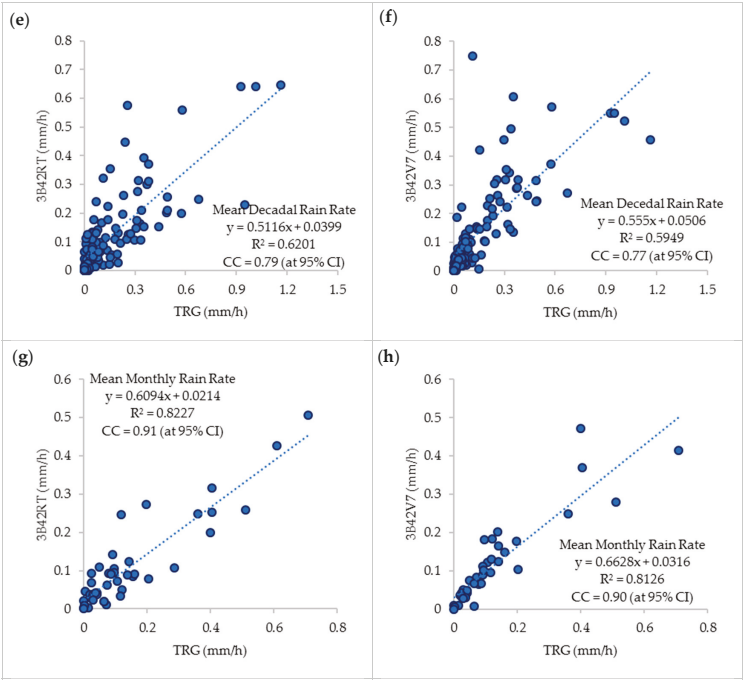


Figure 5. Scatterplots of 3B42RT (left column) and 3B42V7 (right column) rain rates versus TRG rain rates at (**a**,**b**) 3-h interval, (**c**,**d**) mean daily, and (**e**,**f**) mean decadal. The dotted line shows the linear relationship.

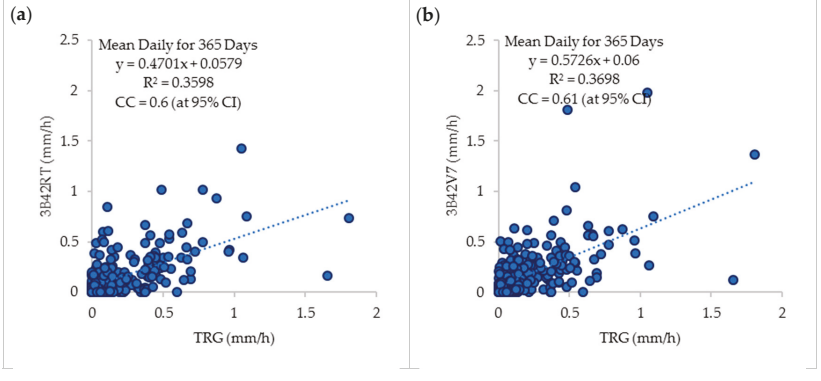


Figure 6. *Cont.*

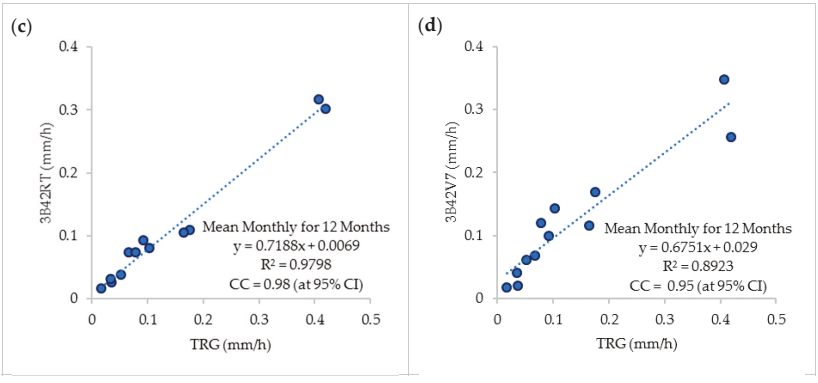


Figure 6. Scatterplots of 3B42RT (left column) and 3B42V7 (right column) rain rates versus TRG rain rates for (**a**,**b**) 365 calendar days and (**c**,**d**) 12 calendar months. The dotted line shows the linear relationship.

Yearly accumulated rainfall comparisons of TRG, 3B43RT, and 3B42v7 are shown in Figure 7. TMPA products underestimate the rainfall amount in all years except 2009, when they slightly overestimate rainfall. Overall, 3B42RT underestimates −24% of the total rainfall amount, while 3B42V7 underestimates −13% of the total rainfall amount.

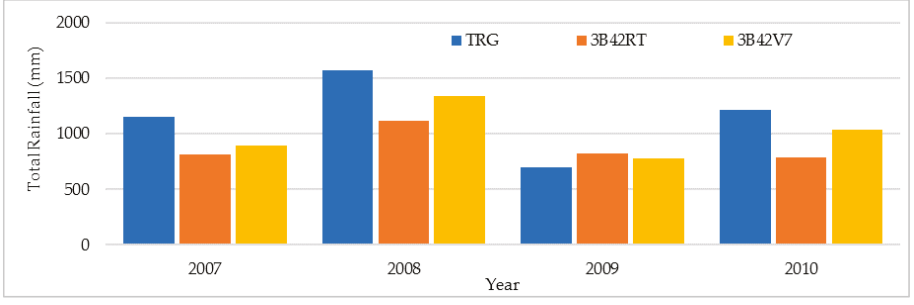


Figure 7. Accumulated rainfall comparison between TMPA and TRG at annual time scales.

3.2. TMPA Products Evaluation for Different Seasons

3.2.1. Seasonal Performance Assessment

Figure 8a,b shows the season-wise various performance indices for 3B42RT and 3B42V7, respectively. 3B42RT performs better in pre-monsoon, with values of 0.73, 0.21, 0.61, and 0.72 for POD, FAR, CSI/TS, and SS, respectively. 3B42V7, on the other hand, has a better performance in post-monsoon, with the highest POD (0.79) and lowest FAR (0.23); TS (0.63) and SS (0.78) are also high compared to other seasons. Post-monsoon is the only season where bias is positive for 3B42V7 (Figure 8c), showing that TMPA detects more rainy days than TRG. 3B42RT has a negative bias for all seasons. The performance of both TMPA products is the lowest during the winter season. The values of POD, FAR, and Bias are (0.24, 0.44), (0.52, 0.34), and (−0.50, −0.33), respectively, for 3B42RT and 3B42V7. Other performance-indices like TS (0.19, 0.36) and SS (0.23, 0.43) for 3B42RT and 3B42V7 are also quite low in comparison to other seasons. It is notable that the winter is typically a dry season in Pakistan, which may explain our present findings. Figure 8 indicates that the performance of both

TMPA products is better in pre-monsoon and post-monsoon, with monsoon results also comparable to these two seasonal periods. However, it is winter, a relatively dry season, when TMPA products perform the worst regarding accurate events detection.

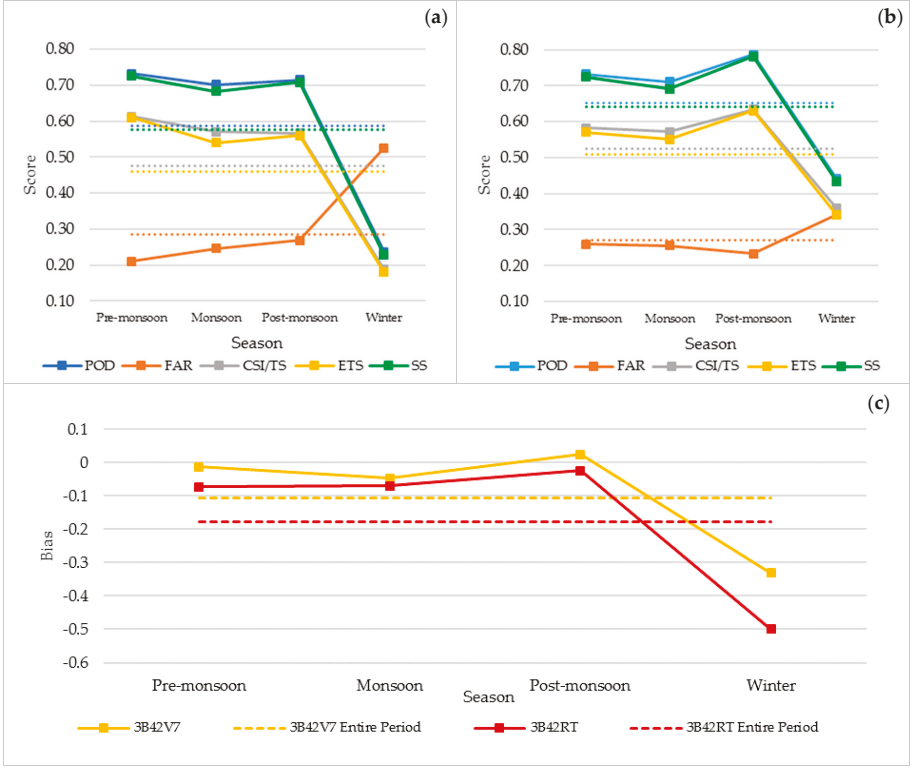


Figure 8. Season-wise various indices score for (**a**) 3B42RT, (**b**) 3B42V7, and (**c**) Bias scores for both TMPA products. Statistics have been calculated for hourly instantaneous rain rates at a 3-h timescale. The dotted lines show scores for the entire period (2007–2010).

3.2.2. TMPA Validation across Different Seasons

Seasonal statistics are summarized in Table 5. The graphical presentation of validation statistics in different seasons is shown in Figure 9.

Table 5. Validation statistics at different time scales in different seasons. Rain rate values at a 3-h interval are instantaneous. Daily and above daily rain rates are averaged over time.

Timescale		Rain Rate (mm/h)			SD (mm/h)			MBE (mm/h)		MAE (mm/h)		RMSE (mm/h)		CC	
		TRG	3B42		TRG	3B42		3B42		3B42		3B42		3B42	
			RT	V7		RT	V7	RT	V7	RT	V7	RT	V7	RT	V7
Pre-monsoon	3-h	1.28	1.02	1.41	2.19	1.59	2.16	−0.26	0.13	1.47	1.58	2.46	2.59	0.19	0.29
	Mean Daily	0.17	0.16	0.22	0.30	0.27	0.39	−0.01	0.05	0.18	0.20	0.32	0.33	0.38	0.58
	Mean Decadal	0.07	0.06	0.08	0.09	0.06	0.11	−0.01	0.01	0.04	0.03	0.06	0.05	0.74	0.95
	Mean Monthly	0.07	0.06	0.09	0.04	0.03	0.06	−0.01	0.02	0.03	0.02	0.04	0.03	0.46	0.92
Monsoon	3-h	3.44	2.00	2.11	6.80	3.57	4.14	−1.44	−1.33	3.28	3.24	6.31	6.54	0.44	0.40
	Mean Daily	0.59	0.43	0.45	0.84	0.64	0.75	−0.16	−0.14	0.48	0.45	0.78	0.80	0.49	0.51
	Mean Decadal	0.34	0.24	0.26	0.30	0.19	0.20	−0.10	−0.08	0.16	0.16	0.23	0.24	0.70	0.65
	Mean Monthly	0.33	0.24	0.26	0.21	0.14	0.14	−0.09	−0.07	0.13	0.10	0.15	0.14	0.85	0.83
Post-monsoon	3-h	2.95	1.67	1.72	5.47	2.51	2.97	−1.28	−1.23	2.66	2.63	5.31	5.17	0.34	0.40
	Mean Daily	0.45	0.29	0.30	0.61	0.43	0.50	−0.16	−0.15	0.38	0.35	0.62	0.59	0.38	0.48
	Mean Decadal	0.12	0.08	0.08	0.12	0.06	0.09	−0.04	−0.04	0.07	0.05	0.11	0.08	0.57	0.83
	Mean Monthly	0.10	0.07	0.07	0.10	0.05	0.07	−0.03	−0.03	0.05	0.03	0.08	0.05	0.59	0.94
Winter	3-h	1.21	0.90	1.20	1.64	2.02	2.10	−0.31	−0.01	1.69	1.64	2.53	2.48	0.07	0.24
	Mean Daily	0.24	0.21	0.29	0.37	0.39	0.44	−0.03	0.05	0.29	0.26	0.48	0.40	0.20	0.54
	Mean Decadal	0.08	0.07	0.09	0.09	0.06	0.10	−0.01	0.02	0.05	0.03	0.08	0.05	0.57	0.85
	Mean Monthly	0.06	0.05	0.07	0.05	0.04	0.06	−0.01	0.01	0.02	0.02	0.03	0.03	0.79	0.94

All seasons follow a similar trend, i.e., correlation improves, and a decrease in errors when transitioning from a short time window (3-h) to a larger time interval average (e.g., monthly). As per Figure 9a,b, 3B42RT underestimates rainfall in all seasons, while 3B42V7 underestimates rainfall in monsoon and post-monsoon seasons; and overestimates rainfall in winter and pre-monsoon at daily, decadal, and monthly rain rate averages. At 3-h time interval, 3B42V7 exhibits a similar pattern, except it slightly underestimates in winter. At a 3-h timescale, a maximum value of MBE (−1.44, −1.33), MAE (3.28, 3.24), and RMSE (6.31, 6.54) has been observed in the monsoon for 3B42RT and 3B42V7, respectively. Post-monsoon is another season where TMPA products highly underestimate three-hourly rain rates (see Table 5). The highest MBE, MAE, and RMSE values at all time intervals for both TMPA products are associated with the monsoon season. Correlation wise, at a 3-h time interval, monsoon and post-monsoon are the seasons where TMPA products have the highest CC values, i.e., 0.44 and 0.34 for 3B42RT, and 0.40, 0.40 for 3B42V7. In the pre-monsoon and winter seasons, CC values are less than 0.3 for both TMPA products. Figure 9e–h shows that the correlation between TMPA products and TRG is higher, while RMSE is reduced when moving from three-hourly towards monthly time intervals. At a 3-h interval, the correlation has the highest values in monsoon and post-monsoon seasons, while the RMSE values are also greater in these two seasons. For daily, decadal, and monthly time intervals, winter and pre-monsoon seasons dominate regarding better CC and lower RMSE for 3B42V7. 3B42RT behaves inconsistently when moving from three-hourly towards the monthly time interval. Standard deviation values across various seasons and timescales show that TMPA products mostly have a lower skill in capturing rainfall variability. The winter season is a specific season where TMPA products are affected more than TRG due to the anomaly in rainfall.

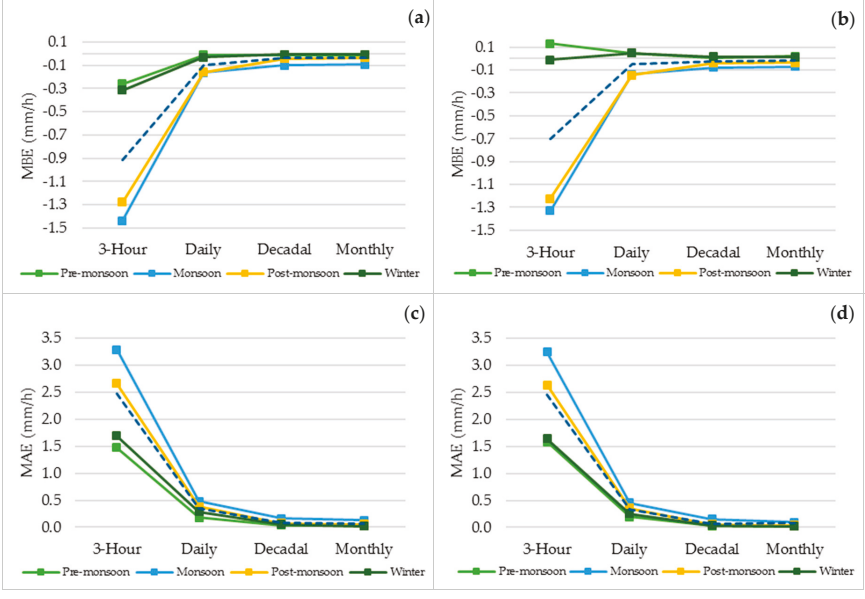


Figure 9. *Cont.*

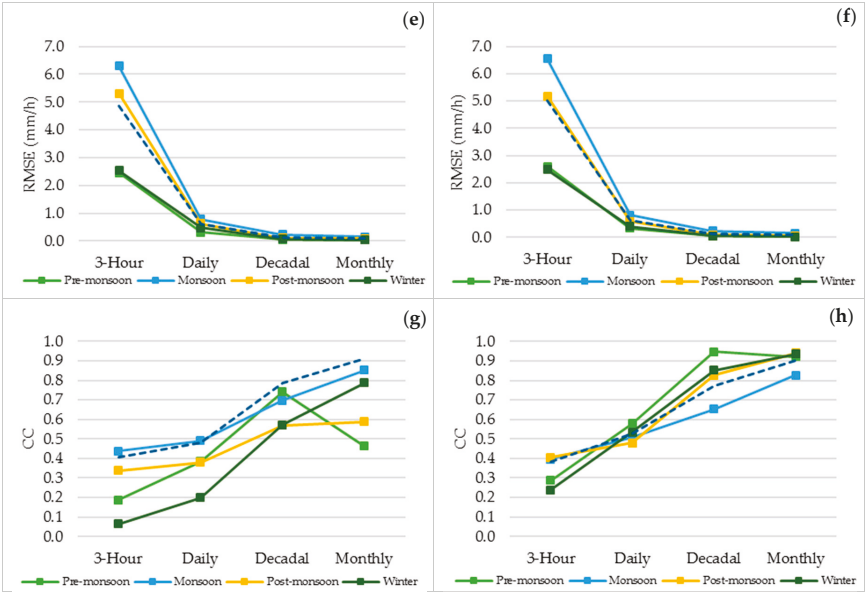


Figure 9. Season-wise validation statistics for 3B42RT (left column) and 3B42V7 (right column) (**a**,**b**) MBE, (**c**,**d**) MAE, (**e**,**f**) RMSE, and (**g**,**h**) CC at different time intervals. The dotted lines show validation scores for the entire time span (2007–2010).

Figure 10 shows the time series of accumulated monthly rainfall. TMPA products and TRG follow a similar rainfall pattern for all years except the 2009 monsoon. In general, TMPA products overestimate in portions of February and April. These two months belong to winter and pre-monsoon seasons. In monsoon months, i.e., June, July, and August, TMPA products largely underestimate rainfall. However, in the 2009 monsoon, 3B42RT overestimates rainfall, which is an exception. The possible reason for this is the very low rainfall intensity and rainfall amount recorded for the monsoon in 2009. In post-monsoon, TMPA products exhibit relatively smaller overestimation. In November and December, rainfall from both measuring sources is quite comparable. The overall trend indicates that TMPA products heavily underestimate in wet periods.

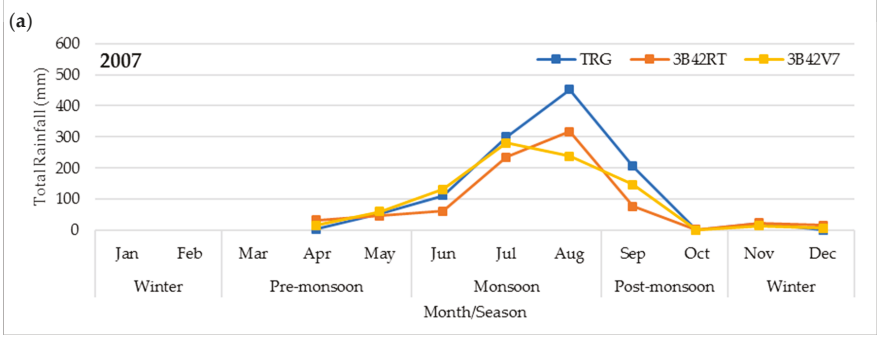


Figure 10. *Cont.*

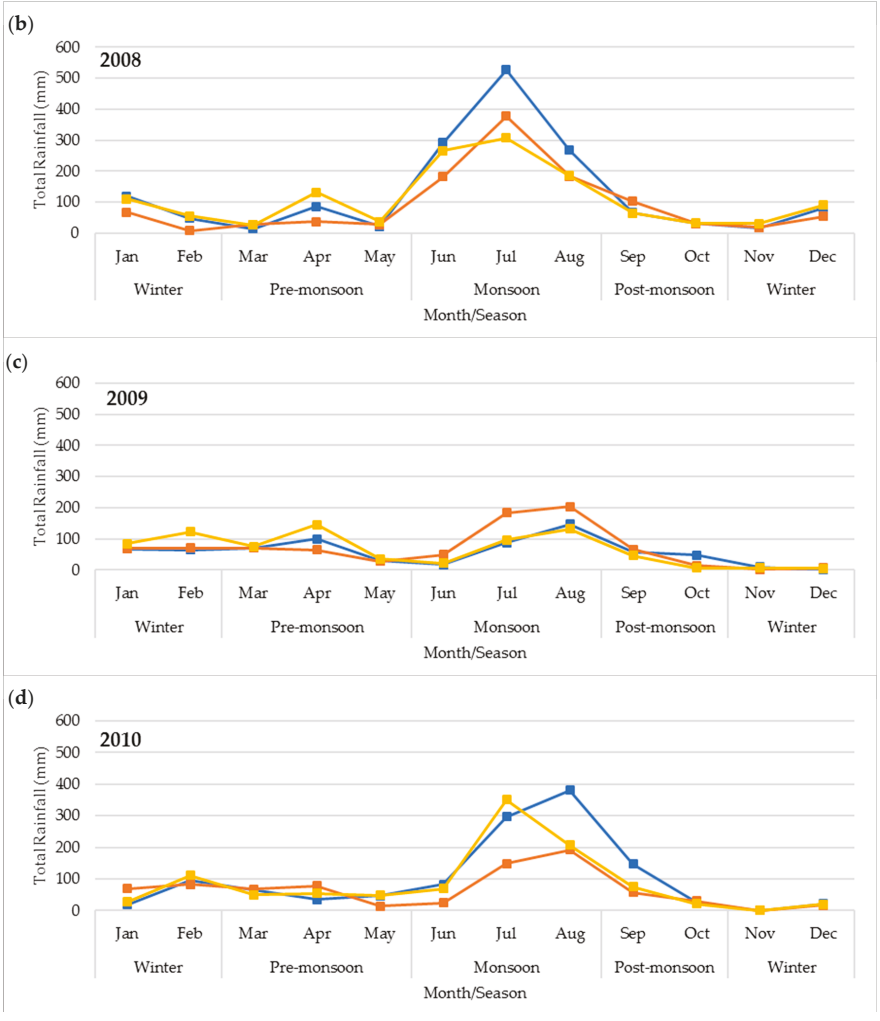


Figure 10. Monthly accumulated rainfall pattern of TMPA and TRG for the (**a**) year 2007, (**b**) year 2008, (**c**) year 2009, and (**d**) year 2010.

Seasonal accumulated rainfall comparisons of TRG, 3B42RT, and 3B42V7 are shown in Figure 11. 3B42RT underestimates rainfall in all seasons (−10% in the winter season, −7% in the pre-monsoon season, −27% in the monsoon season, and −35% in the post-monsoon season). Contrarily, 3B42V7 overestimates rainfall in the winter (21%) and pre-monsoon (29%) seasons, while underestimates in the monsoon (−23%) and post-monsoon (−32%) seasons. Here, a negative sign indicates underestimation.

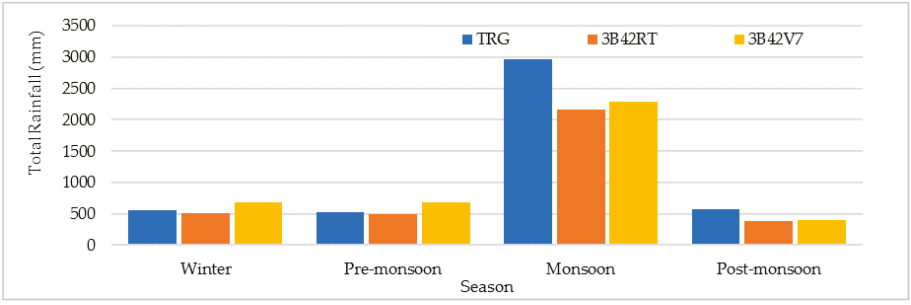


Figure 11. Accumulated rainfall comparison between TMPA and TRG at the seasonal timescale.

3.3. Diurnal Rainfall Analysis

3.3.1. Diurnal Rainfall Analysis for Entire Study Period

The diurnal patterns in terms of average rainfall intensity, total rainfall amount, and frequency of rain occurrence over the entire period are shown in Figure 12. The results indicate that the TRG has higher rainfall intensities throughout the day, except for the evening (17:00 h local), where 3B42V7 has a slightly greater mean rainfall intensity. In terms of rainfall amount, both TMPA products overestimate around midnight, while 3B42V7 overestimates in the evening as well. The intensity curve of TRG clearly shows two peaks, one at midnight and another one in the afternoon. However, rainfall amount has a single strong peak in the morning (05:00 h local). A comparative analysis of Figure 12 reveals that Lai Nullah basin received a smaller number of rainy events in the evening, but these were intensive showers. Maximum rainfall amount was received between midnight and the morning, where rainfall intensities vary from high to low. TMPA products complement each other quite well, except for evening hours, but their agreement with TRG is quite weak. At a 02:00 h local time, TRG and TMPA products have similar frequencies, with TRG having higher rainfall intensities. But at this specific hour, TMPA rainfall amount exceeds TRG rainfall amount, which is quite unconventional. One of the basic reasons behind this exception is the difference in the calculation mechanism for total rainfall. TRG considers rain rates at each hour of the day to calculate the total three-hourly rainfall amount. TMPA products, on the other hand, have only one value of hourly rain rates for each 3-h interval. Consequently, hourly rain rates at each 3-h interval are multiplied by '3' to get the three-hourly rainfall amount. The difference in rainfall amounts signifies a greater variation associated with TRG rainfall intensities between a 23:00 to 02:00 h local time, which TMPA products are unable to detect.

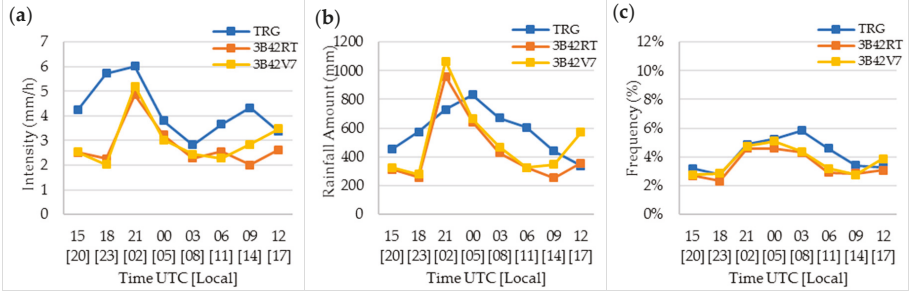


Figure 12. Diurnal cycle of precipitation: (**a**) average rainfall intensity, (**b**) total rainfall amount, and (**c**) frequency of rainfall occurrence.

3.3.2. Seasonal Diurnal Rainfall Analysis

The season-wise diurnal cycle is depicted in Figure 13. In pre-monsoon seasons, TMPA products, especially 3B42V7, follow TRG curve patterns. The TMPA products underestimate rainfall in the morning and overestimate in the evening. The TRG has higher rainfall intensities and percentage frequency at midnight (02:00 h local time). TMPA products have higher percentage frequencies and rainfall intensities during evening hours.

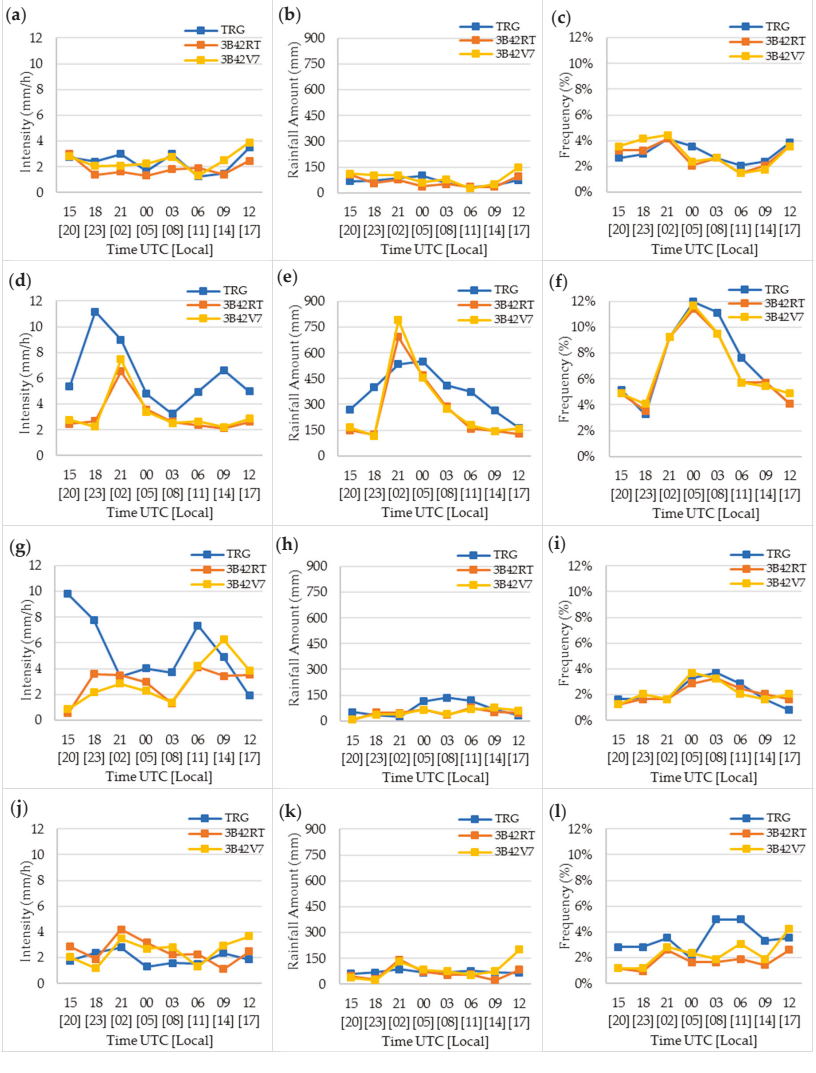


Figure 13. Season-wise diurnal precipitation patterns for: (**a**–**c**) pre-monsoon, (**d**–**f**) monsoon, (**g**–**i**) post-monsoon and (**j**–**l**) winter seasons. Rainfall intensities are in the left column, the rainfall amount is in the middle, and rainfall frequency is in the right column.

As discussed earlier, the monsoon season is the wet season in Pakistan, and almost 60% of annual rainfall amount is received in this season. As a result, the monsoon season has a greater impact on

rainfall patterns calculated for this entire period. Figure 10d,e displays the diurnal monsoon patterns, which are very similar to the overall time series patterns. It is notable that TRG has a higher percentage frequency during the daytime. Overall, the TMPA products underestimate rainfall, except for a single time slot, while their intensities are lower than TRG intensities. The rainfall intensity pattern in post-monsoon is quite dispersed. Even TMPA products mutually do not complement each other. TMPA products underestimate rainfall in morning and afternoon hours. The diurnal cycle of the winter season presents an altogether different trend, where TMPA products generally have higher rainfall intensities but lower percentage frequencies. The rainfall amount variation between 3B42V7 and TRG is greater in the evening, i.e., 17:00 h local time.

3.3.3. Case Study: Heavy Rainfall Event—July 2008

Afzal et al. [14] reported heavy rain events in Lai Nullah (current study area) during July 2008. Three-hourly rain rates (mm/h) for both TMPA products have been examined with TRG observations (Figure 14). Out of 48 events in six days, TRG detects ten rainy days, while both TMPA products sense 12 rainy days. TMPA products miss one rainy day and give three false alarms about TRG observations. The maximum rain rate observed by TRG is 18.8 mm/h on 5th July (6:00 h UTC, 11:00 h local), while the 3B43RT and 3B42V7 rain rate at that specific time slot is 19.1 mm/h and 15.8 mm/h, respectively.

The total rainfall amount for the event is shown in Figure 15. TMPA rainfall products follow a pattern similar to the TRG rainfall, but both products severely underestimated the total rain amount. In spite of monthly adjustments, 3B42V7 only estimated 52% of total rainfall. 3B42RT, on the other hand, estimated 64% of total rainfall.

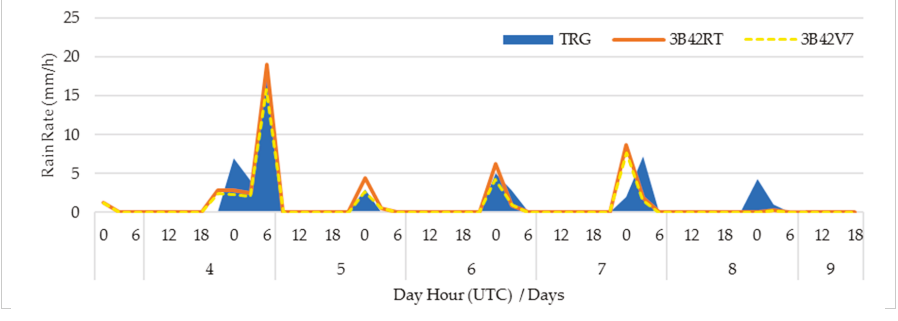


Figure 14. Heavy rain event in July 2008: (a) rain rate (mm/h) pattern for TRG, 3B42RT, and 3B42V7 data at the three-hourly time interval.

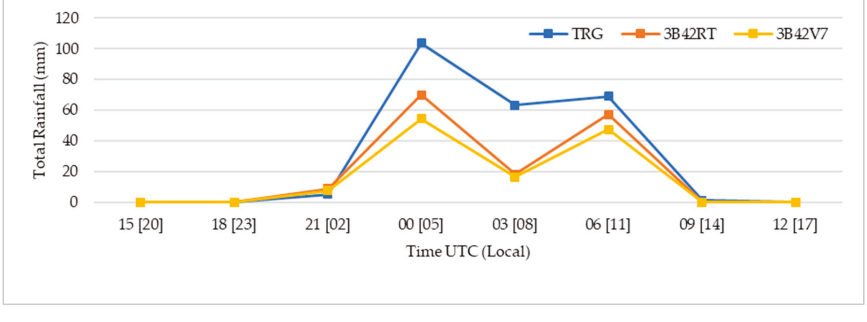


Figure 15. Total rainfall amount calculated from 4 July 2008 to 9 July 2008.

Daily patterns of rainfall over the regional scale, starting from 1st of July 2008, are shown in Figure 16. These patterns are analyzed by using TMPAV7 daily rainfall data. The spatial rainfall patterns indicate that July 2008 heavy rainfall events in Lai Nullah basin were part of a Monsoon system that entered the north-east side of Pakistan from India. On 1 July 2008, a strong Monsoon system hit south-west India, and it propagated towards north-eastern parts of India. Finally, the system impacted the great Himalaya mountains and deflected along the Himalaya's foothills towards Pakistan. It entered Pakistan on the 3rd of July and caused heavy rainfall events in Lai Nullah basin on the 4th of July and afterwards.

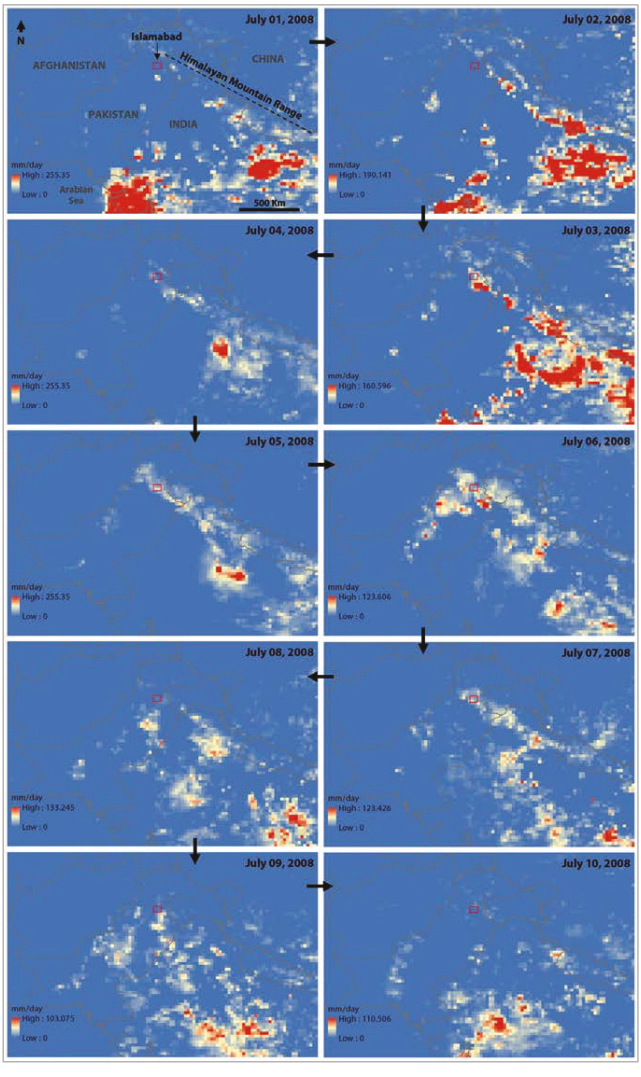


Figure 16. Daily rainfall patterns at the regional scale; from 1 July 2008 to 10 July 2008. Red box approximates study area location.

4. Discussion

Our study evaluates TMPA products at various temporal scales, including the best-offered one, i.e., the 3-h timescale. Overall, both of the TMPA products underestimated rainfall across the entire study period. However, 3B72V7 performed better than 3B42RT, which agrees with the finding of various recently conducted studies [53,54,57,77,78] in Pakistan. Further, our study strengthens the fact that the performance of TMPA products depends upon seasons and various timescales [16,79]. Categorical indices like bias, POD, FAR, CSI/TS, and SS values for 3B42V7 are −0.11, 0.65, 0.27, 0.52, 0.51, and 0.64, respectively. These scores are much better when compared to two erstwhile studies [48,52]. Categorical indices' scores at different rainfall intensities reveal that the TMPA products detect plenty of spurious rain events at light rain rates (rain rates <2 mm/h). The value of FAR for light rain events is 0.76 for 3B42RT and 0.71 for 3B42V7. Figure 3a,b show that all fallacious alarms befall at rain intensities below 2 mm/h, which is concurrent with the findings of [47]. A possible reason for the high FAR at low rain rates could be wind induced errors during rain gauge observations. A score of FAR ameliorates and attains its ideal value when shifted from low rainfall intensities towards higher rainfall intensities by keeping the rain detection threshold constant, i.e., 0.1mm. TMPA's performance in the categorization of various rainfall intensities has been examined in Figure 4. Results show that the TMPA products behave haphazardly and misclassify many 'light rain' events into 'moderate' and 'heavy rain' events and vice versa. Low standard deviation values for TMPA products can explain the inability of these products to detect variations in rainfall intensities. Overall, 3B42V7's performance is better than 3B42RT in rain intensities classification.

Rainfall rates validation results for the entire study period (Table 4) show a very small value of CC = 0.41 for 3B42RT and CC = 0.38 for 3B42V7 at a three-hourly timescale for the entire time span. The CC value for 3B42V7 is smaller than the value calculated in [52] that gives an average value of CC = 0.45. The calculated MBE is −0.92 and −0.70 for 3B42RT and 3B42V7, respectively, which indicates an underestimation of TMPA products in rainfall measurements. Standard deviation values show that both TMPA products detect small variability in rainfall compared to reference rainfall data. However, 3B42V7 detects better variation in rain occurrence compared to 3B42RT. Our results second what Anjum et al. [54] concluded in their study, i.e., the inability of TMPA products to capture temporal variability. At daily, decadal, and monthly time scales, the correlation values between TMPA products and TRG rain rates increased to (0.48, 0.53), (0.79, 0.77), and (0.91, 0.90), respectively, which are compatible with the results of a couple of bygone studies [16,47]. However, CC values at a daily timescale are less than the values (0.57, 0.61) calculated by Anjum et al. [78]. The cumulative rainfall analysis shows an overall underestimation of TMPA products in measuring actual rainfall. By considering the standard definition of acceptable accuracy, i.e., ±10% of observed precipitation [43,44], 3B42RT exhibits more underestimation of rainfall (−24%) than 3B42V7 (−13%). Our result is aligned with the findings of Anjum et al. [54].

Figure 8 shows that the TMPA products' performance is congruous in pre-monsoon, monsoon, and post-monsoon seasons. Nonetheless, these products have the poorest (or lowest) performance in the winter season, where the score for FAR is as high as 0.52 (for 3B42RT) and scores of all other indices are less than what is calculated in other seasons. Further, the value of the correlation coefficient is very low in winter for both TMPA products and these products overestimate rainfall in the winter season. Winter is a relatively dry season (occasional rainfall events and low rainfall intensities) in Pakistan. Therefore, the performance of the TMPA products in the dry season is insignificant, which is congruent to the findings of [47]. One possible reason for rainfall overestimation and high false alarms in winter is that raindrops may evaporate before reaching the surface in dry conditions. One more possibility is a higher concentration of atmospheric aerosols in the winter season, which may intrude the rainfall and subsequently diminish the precipitation amount observed by a surface rain gauge. A higher aerosol concentration, on the contrary, can increase the cloud drop size as well, which may eventually cause overestimation by a satellite sensor [80]. Another aspect of low agreement between satellite-based rainfall measurements and rain gauge observations in winter or at light rainfall intensities is the

errors associated with in-situ rain gauge observations. Tipping bucket (TB) rain gauges may include both systematic as well as random errors at light rainfall intensities. Systematic error may include calibration error, wind-induced error, and sampling error (especially at high temporal sampling) [81]. Thus, at low rainfall intensities, uncertainties are not only associated with satellite measured rainfall, but may also lay with rain gauge observations. Error statistics show that MBE and RMSE have greater values in the monsoon season across all time scales, which agrees with the fact that TMPA maintains greater disparities in the wet season or at high rainfall intensities [43–45,47]. Both of the products largely underestimate rainfall in the monsoon season. High-intensity rain events are associated with the monsoon season, and such heavy rainfall intensities may cause signal attenuation of a passive microwave sensor, leading to underestimation in satellite-based precipitation products [80]. Figure 9 gives a good insight into the behavior of CC and RMSE between TMPA products' rain rates and TRG rain rates at various temporal scales. It should be noted that 3B42RT has an inconsistent behavior at different time scales.

The diurnal rainfall analysis indicates an inconsistent pattern of TMPA products with respect to TRG diurnal pattern, which supports the results of a similar study [49] conducted in China. TMPA products are quite aligned in the pre-monsoon season, but behave differently in other seasons. Monsoon is the most crucial season in Pakistan as it receives a major portion of annual rainfall amount. In this season, the county faces a high frequency of floods. Lai Nullah also receives flash floods typically in the monsoon season. Monsoon diurnal analysis suggests that possible flooding hours are from 02:00 to 11:00 h local time. In these time slots, TRGs show the maximum rainfall amount, with a peak at 05:00 h local time. Also, the rainfall intensities in these time slots have greater variations. Thus, intensive rainfall intensities in these time slots may produce flash floods in Lai Nullah basin. But TMPA products, on the other hand, underestimate rainfall in identifiable time periods, which questions their ability to detect flash floods. This discrepancy has been further emphasized by the July 2008 heavy rainfall events case study. The spatial pattern shows that the July 2008 event was part of the monsoon system generated from the north-eastern directions (India and Nepal) of the country. The case study has shown that TMPA products severely underestimated rainfall in this short-duration hydrometeorological event. Similar results are presented by Anjum et al. during their study [54] on extreme heavy precipitation events over Swat River Basin. In our study, 3B42RT underestimated 36% of the total rainfall amount, while3B42V7, which is a monthly adjusted product, performed even worse and underestimated 48% percent of the total rainfall. Moreover, TMPA products are also not satisfactory in post-monsoon and winter seasons in terms of their diurnal performance.

5. Conclusions

The suitability of the TMPA three-hourly products is studied at the flash flooding site of Lai Nullah basin in Islamabad, Pakistan, using 10-min telemetric rain gauge data. Urban flash flooding is a localized and heterogeneous hydro-meteorological phenomenon and to predict such phenomena, a high spatio-temporal rainfall data in real or near-real time is crucial. The prediction of urban flash flooding is perplexing due to the dearth of good quality in-situ rainfall data. Further, many hydrological models need rainfall at high spatial and temporal resolutions. Satellite sensor-based products may prove to be a good substitute to in-situ rainfall data given that they have a reasonable accuracy. Accurate satellite-based rainfall datasets may assist the scientific community to eliminate rainfall data gaps issues. The current study examines both TMPA products, i.e., a near-real time product (3B42RT) and research version product (3B42V7). Both TMPA products are evaluated at their best possible temporal resolution, i.e., three-hourly rain rates, to check their suitability for flash flood early warnings and water-related modeling. In the current study, scores of categorical indices (POD, FAR, CSI, and SS) show that the performance of TMPA research version in the detection of rain/no-rain events is quite good and has the best scores in a post-monsoon season. However, the performance of the TMPA near-real time product is inferior to the research version. Both TMPA products perform least in the winter season and have the lowest scores for performance indices. Furthermore, the performance of

TMPA products is questionable at light rain rates, but it improves at higher rain intensities. Regarding continuous statistical measures, the current study reveals that the correlation coefficient between TMPA (3B42RT, 3B42V7) estimates and telemetric rain gauge observations increases from a modest value (0.41, 0.38) at a three-hourly time interval to a high value (0.91, 0.90) at monthly averages. Moreover, higher mean absolute error and root mean square error are associated with TMPA at three-hourly rain rate measurement. Nevertheless, these errors are reduced at daily, decadal, and monthly rain rate averages. From a seasonal perspective, the TMPA research version underestimates rainfall in monsoon (wet period) and post-monsoon (relatively less wet period), while overestimates rainfall in winter and pre-monsoon (relatively dry seasons). The TMPA near-real time product continuedly underestimates rainfall over all the seasons. Moreover, monsoon and post-monsoon seasons have larger error values compared to winter and pre-monsoon seasons for both products. Values of the correlation coefficient suggest that TMPA estimates are not a substitute of rain gauge data at 3-h and daily time intervals in Pakistan. Diurnal analysis supported by a heavy rainfall case study suggests that the TMPA near-real time product alone is not a good option to predict a localized phenomenon like urban flash flooding in Pakistan. Both TMPA products have a low linear correlation with reference to in-situ rainfall data and higher bias errors at a three-hourly timescale. However, these products can be utilized in synergy with other rainfall datasets. Overall, both TMPA products are more useful at decadal and above decadal time scales, where these products show a significant value for the correlation coefficient (>0.7) with in-situ reference data and exhibit tolerable errors. In comparison, the TMPA research version shows more consistent behavior at various time scales, while the TMPA near-real-time version has severe issues with its consistency at different time scales. Thus, in circumstances when valid ground gauge data is available, TMPA data (near-real-time version in particular) below a decadal scale should be avoided in Pakistan, particularly in hilly/semi-hilly areas where a high probability of errors is associated with TMPA products due to complex topography. Nonetheless, for the extreme conditions like an area with very sparse rain gauge coverage or none at all, both TMPA rainfall products can be a beneficial source at daily and above daily scales. It is suggested to apply bias correction on both products when using them at daily or below daily timescales. Usage of the TMPA 3B42 near-real time product in critical applications like a land slide and flash floods early warning system, or water-related modeling, is not recommended due to higher errors associated with it, and its inconsistent linear relationship with in-situ rainfall data at various timescales. The TMPA research version, however, at a daily and more reliably, above daily time scale, is a good choice for understanding climate patterns, and it can also be used in applications like water-related modeling. Based on these findings, we intend to continue our work of evaluating satellite precipitation products over this flash flooding site with the GPM IMERG and CMORPH products over suitable periods.

Author Contributions: This study is part of a graduate thesis dissertation project by A.U.R., supervised by F.C. A.U.R. and F.C. designed the entire study framework. A.U.R. performed the analysis, compiled results, and co-wrote the manuscript under the guidance of F.C. W.A.Q. and S.G. supported in MATLAB scripting in preprocessing and data pairing of TMPA measurements and in telemetric rain gauge observations. I.S. provided a critical review of the study methodology and provided insights which helped further analysis of findings. K.F. helped in the compilation of results and manuscript writing.

Funding: This research received no external funding.

Acknowledgments: Authors would like to acknowledge the Pakistan Meteorological Department (PMD) for providing telemetric rain gauge data at a ten-minute timescale for Lai nullah basin Islamabad, Pakistan. Authors also acknowledge the National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA) for making TRMM data freely available at https://storm.pps.eosdis.nasa.gov/storm/. We also thank the anonymous referees for constructive suggestions which improved the quality of this work.

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

References

1. 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]
2. Nasreen, Z.; Zeeshan, M. Environmental hazards faced by inhabitants of Nala Lai. *Pak. Assoc. Anthropol.* **2015**, *27*, 781–784.
3. Shamim, N. Climate change, flood disasters, human displacement and changing lives and food patterns of marginalized people living in tributeries of Nullah Lai Rawalpindi. *Explor. Islam. J. Soc. Sci.* **2016**, *2*, 99–102.
4. Hayat, A. Flash flood forecasting system for Lai Nullah basin (a case study of March 19, 2007 rainfall event). *Pak. J. Meteorol.* **2007**, *4*, 75–84.
5. Ahmad, F.; Kazmi, S.F.; Pervez, T. Human response to hydro-meteorological disasters: A case study of the 2010 flash floods in Pakistan. *J. Geogr. Reg. Plan.* **2011**, *4*, 518–524.
6. Butt, M. A view from Pakistan: Recurrent flash floods in areas surrounding Nullah Lai, Rawalpindi, Pakistan. *Middle East J. Bus.* **2014**, *9*, 29–32. [CrossRef]
7. Hashmi, H.N.; Siddiqui, Q.T.M.; Ghumman, A.R.; Kamal, M.A.; Mughal, H.R. A critical analysis of 2010 floods in Pakistan. *Afr. J. Agric. Res.* **2012**, *7*, 1054–1067.
8. Paulikas, M.J.; Rahman, M.K. A temporal assessment of flooding fatalities in Pakistan (1950–2012). *J. Flood Risk Manag.* **2015**, *8*, 62–70. [CrossRef]
9. Wikipedia Lai Nullah. 2017. Available online: https://en.wikipedia.org/wiki/Lai_Nullah (accessed on 21 January 2018).
10. Ahmad, B.; Kaleem, M.S.; Butt, M.J.; Dahri, Z.H. Hydrological modelling and flood hazard mapping of Nullah Lai. *Proc. Pak. Acad. Sci.* **2010**, *47*, 215–226.
11. Kamal, A. *Pakistan: Lai Nullah Basin Flood Problem Islamabad—Rawalpindi Cities*; World Meteorological Organization: Geneve, Switzerland, 2004.
12. Japan International Cooperation Agency. *Basic Design Study Report on the Lai Nullah Flood Forecasting and Warning System Project in the Islamic Republic of Pakistan*; Japan International Cooperation Agency: Tokyo, Japan, 2005.
13. Siddiqui, Q.T.M.; Hashmi, H.N.; Ghumman, A.R.; Mughal, H.R. Flood Inundation Modeling for a Watershed in the Pothowar Region of Pakistan. *Arab. J. Sci. Eng.* **2011**, *36*, 1203–1220. [CrossRef]
14. Afzal, M.; Zaman, Q. Case study: Heavy rainfall event over Lai Nullah catchment area. *Pak. J. Meteorol.* **2010**, *6*, 39–48.
15. Rehman, H.; Chaudhry, A.G.; Nasir, A. Nala Lai and urban ecology: A study of dwellers and disaster management. *Pak. Assoc. Anthropol.* **2015**, *27*, 617–619.
16. Iqbal, M.F.; Hussain, A. Validation of satellite based precipitation over diverse topography of Pakistan. *Atmos. Res.* **2018**, *201*, 247–260. [CrossRef]
17. Siuki, S.K.; 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]
18. Boegh, E.; Thorsen, M.; Butts, M.B.; Hansen, S.; Christiansen, J.S.; Abrahamsen, P.; Hasager, C.B.; Jensen, N.O.; van der Keur, P.; Refsgaard, J.C.; et al. Incorporating remote sensing data in physically based distributed agro-hydrological modelling. *J. Hydrol.* **2004**, *287*, 279–299. [CrossRef]
19. Huffman, G.J. Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.* **2001**, *2*, 36–50. [CrossRef]
20. Ebert, E.E. Methods for verifying satellite precipitation estimates. In *Measuring Precipitation from Space: EURAINSAT and the Future*; Levizzani, V., Bauer, P., Turk, F.J., Eds.; Springer: Dordrecht, The Netherlands, 2007; Volume 28, pp. 345–356.
21. 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]
22. Joyce, R.J.; Janowiak, J.E.; Arkin, 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]

23. Hsu, K.-L.; 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]
24. Rauniyar, S.P.; Protat, A.; Kanamori, H. Uncertainties in TRMM-Era multisatellite-based tropical rainfall estimates over the Maritime Continent. *Earth Space Sci.* **2017**, *4*, 275–302. [CrossRef]
25. Worqlul, A.W.; Yen, H.; Collick, A.S.; Tilahun, S.A.; Langan, S.; Steenhuis, T.S. Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions: The upper Blue Nile Basin, Ethiopia. *Catena* **2017**, *152*, 242–251. [CrossRef]
26. Roh, W.; Satoh, M.; Nasuno, T. Improvement of a cloud microphysics scheme for a global nonhydrostatic model using trmm and a satellite simulator. *J. Atmos. Sci.* **2017**, *74*, 167–184. [CrossRef]
27. Hong, Y.; Adler, R.F.; Negri, A.; Huffman, G.J. Flood and landslide applications of near real-time satellite rainfall products. *Nat. Hazards* **2007**, *43*, 285–294. [CrossRef]
28. Yoshimot, S.; Amarnath, G. Applications of satellite-based rainfall estimates in flood inundation modeling—A case study in Mundeni Aru River Basin, Sri Lanka. *Remote Sens.* **2017**, *9*, 998. [CrossRef]
29. Ricko, M.; Carton, J.A.; Birkett, C. Climatic effects on lake basins. Part I: Modeling tropical lake levels. *J. Clim.* **2011**, *24*, 2983–2999. [CrossRef]
30. Wu, H.; Adler, R.F.; Hong, Y.; Tian, Y.; Policelli, F. Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. *J. Hydrometeorol.* **2012**, *13*, 1268–1284. [CrossRef]
31. Kirschbaum, D.; Adler, R.; Adler, D.; Peters-Lidard, C.; Huffman, G. Global distribution of extreme precipitation and high-impact landslides in 2010 relative to previous years. *J. Hydrometeorol.* **2012**, *13*, 1536–1551. [CrossRef]
32. Liu, X.; Liu, F.M.; Wang, X.X.; Li, X.D.; Fan, Y.Y.; Cai, S.X.; Ao, T.Q. Combining rainfall data from rain gauges and TRMM in hydrological modelling of Laotian data-sparse basins. *Appl. Water Sci.* **2017**, *7*, 1487–1496. [CrossRef]
33. Froidurot, S.; Diedhiou, A. Characteristics of wet and dry spells in the West African monsoon system. *Atmos. Sci. Lett.* **2017**, *18*, 125–131. [CrossRef]
34. De Jesús, A.; Breña-Naranjo, J.A.; Pedrozo-Acuña, A.; Yamanaka, V.H.A. The use of TRMM 3B42 product for drought monitoring in Mexico. *Water (Switzerland)* **2016**, *8*, 325. [CrossRef]
35. Yong, B.; Liu, D. 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]
36. 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]
37. Wu, H.; Adler, R.F.; Tian, Y.; Huffman, G.J.; Li, H.; Wang, J. Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model. *Water Resour. Res.* **2014**, *50*, 2693–2717. [CrossRef]
38. Chang, L.T.-C.; McAneney, J.; Cheung, K.K.W. Case study of trmm satellite rainfall estimation for landfalling tropical cyclones: Issues and challenges. *Trop. Cyclone Res. Rev.* **2013**, *2*, 109–123.
39. Prakash, S.; Mahesh, C.; Gairola, R.M.; Pal, P.K. Comparison of high-resolution TRMM-based precipitation products during tropical cyclones in the North Indian Ocean. *Nat. Hazards* **2012**, *61*, 689–701. [CrossRef]
40. National Space Development Agency. *TRMM Data Users Handbook*; National Space Development Agency of Japan: Saitama, Japan, 2001.
41. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116. [CrossRef]
42. Moazami, S.; Golian, S.; Kavianpour, M.R.; Hong, Y. Comparison of PERSIANN and V7 TRMM Multi-satellite Precipitation Analysis (TMPA) products with rain gauge data over Iran. *Int. J. Remote Sens.* **2013**, *34*, 8156–8171. [CrossRef]
43. Almazroui, M. Calibration of TRMM rainfall climatology over Saudi Arabia during 1998–2009. *Atmos. Res.* **2010**, *99*, 400–414. [CrossRef]
44. Islam, M.N.; Das, S.; Uyeda, H. Calibration of TRMM derived rainfall over Nepal during 1998–2007. *Open Atmos. Sci. J.* **2010**, *4*, 12–23. [CrossRef]
45. Islam, M.N.; Uyeda, H. Use of TRMM in determining the climatic characteristics of rainfall over Bangladesh. *Remote Sens. Environ.* **2007**, *108*, 264–276. [CrossRef]

46. Cai, Y.; Jin, C.; Wang, A.; Guan, D.; Wu, J.; Yuan, F.; Xu, L. Comprehensive precipitation evaluation of TRMM 3B42 with dense rain gauge networks in a mid-latitude basin, northeast, China. *Theor. Appl. Climatol.* **2016**, *126*, 659–671. [CrossRef]
47. Scheel, M.L.M.; Rohrer, M.; Huggel, C.; Villar, D.S.; Silvestre, E.; Huffman, G.J. Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 2649–2663. [CrossRef]
48. Kotsıfakıs, K.; Felonı, E.; Kotronı, V.; Baltas, E. Evaluation of the Version 7 TRMM Multi-Satellite Precipitation Analysis (TMPA) 3B42 product over Greece. In Proceedings of the 15th International Conference on Environmental Science and Technology, Rhodes, Greece, 31 August–2 September 2017.
49. Yuan, W.; Li, J.; Chen, H.; Yu, R. Intercomparison of summer rainfall diurnal features between station rain gauge data and TRMM 3B42 product over central eastern China. *Int. J. Climatol.* **2012**, *32*, 1690–1696. [CrossRef]
50. Sohn, B.J.; Han, H.J.; Seo, E.K. Validation of satellite-based high-resolution rainfall products over the Korean Peninsula using data from a dense rain gauge network. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 701–714. [CrossRef]
51. Villarini, G. Evaluation of the research-version TMPA rainfall estimate at its finest spatial and temporal scales over the Rome metropolitan area. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 2591–2602. [CrossRef]
52. 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]
53. Hussain, Y.; Sathe, F.; Hussain, M.B.; Martinez, H.E. Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theor. Appl. Climatol.* **2017**, *131*, 1–14. [CrossRef]
54. Anjum, M.N.; Ding, Y.; Shangguan, D.; Ijaz, M.W.; Zhang, S. Evaluation of high-resolution satellite-based real-time and post-real-time precipitation estimates during 2010 extreme flood event in Swat River Basin, Hindukush Region. *Adv. Meteorol.* **2016**, *2016*, 1–8. [CrossRef]
55. 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. *Atmos. Sci. Lett.* **2016**, *17*, 270–279. [CrossRef]
56. Khan, S.I.; Hong, Y.; Gourley, J.J.; Huelsing, H.K.; Khattak, M.U.; Vergara, H.J. Spatial and diurnal variability of monsoon systems assessed by TRMM rain rate over Indus Basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4325–4335. [CrossRef]
57. 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]
58. Ali, M.; Khan, S.J.; Aslam, I.; Khan, Z. Simulation of the impacts of land-use change on surface runoff of Lai Nullah Basin in Islamabad, Pakistan. *Landsc. Urban Plan.* **2011**, *102*, 271–279. [CrossRef]
59. Jarraud, M. Measurement of precipitation. In *Guide to Meteorological Instruments and Methods of Observation (WMO-No. 8)*, 7th ed.; World Meteorological Organisation: Geneva, Switzerland, 2008.
60. Liu, Z. Comparison of versions 6 and 7 3-hourly TRMM multi-satellite precipitation analysis (TMPA) research products. *Atmos. Res.* **2015**, *163*, 91–101. [CrossRef]
61. 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: New York, NY, USA, 2010; pp. 3–22.
62. Prakash, S.; Gairola, R.M. Validation of TRMM-3B42 precipitation product over the tropical Indian Ocean using rain gauge data from the RAMA buoy array. *Theor. Appl. Climatol.* **2014**, *115*, 451–460. [CrossRef]
63. Porcù, F.; Milani, L.; Petracca, M. On the uncertainties in validating satellite instantaneous rainfall estimates with raingauge operational network. *Atmos. Res.* **2014**, *144*, 73–81. [CrossRef]
64. Yu, S.; Eder, B.; Dennis, R.; Chu, S.-H.; Schwartz, S.E. New unbiased symmetric metrics for evaluation of air quality models. *Atmos. Sci. Lett.* **2006**, *7*, 26–34. [CrossRef]
65. Tang, G.; Zeng, Z.; Long, D.; Guo, X. 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]

66. Kumar, D.; Gautam, A.K.; Palmate, S.S.; Pandey, A.; Suryavanshi, S.; Rathore, N.; Sharma, N. Evaluation of TRMM multi-satellite precipitation analysis (TMPA) against terrestrial measurement over a humid sub-tropical basin, India. *Theor. Appl. Climatol.* **2017**, *129*, 783–799. [CrossRef]
67. 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. *Theor. Appl. Climatol.* **2018**, *131*, 377–398. [CrossRef]
68. Yin, S.; Chen, D.; Xie, Y. Diurnal variations of precipitation during the warm season over China. *Int. J. Climatol.* **2009**, *29*, 1154–1170. [CrossRef]
69. Jeong, J.H.; Walther, A.; Nikulin, G.; Chen, D.; Jones, C. Diurnal cycle of precipitation amount and frequency in Sweden: Observation versus model simulation. *Tellus Ser. A Dyn. Meteorol. Oceanogr.* **2011**, *63*, 664–674. [CrossRef]
70. Pfeifroth, U.; Trentmann, J.; Fink, A.H.; Ahrens, B. Evaluating satellite-based diurnal cycles of precipitation in the African tropics. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 23–39. [CrossRef]
71. Haider, K.; Khokhar, M.F.; Chishtie, F.; RazzaqKhan, W.; Hakeem, K.R. Identification and future description of warming signatures over Pakistan with special emphasis on evolution of CO_2 levels and temperature during the first decade of the twenty-first century. *Environ. Sci. Pollut. Res.* **2017**, *24*, 7617–7629. [CrossRef]
72. Hussain, S.; Xianfang, S.; Hussain, I.; Jianrong, L.; Mei, H.D.; Hu, Y.L.; Huang, W. Controlling factors of the stable isotope composition in the precipitation of Islamabad, Pakistan. *Adv. Meteorol.* **2015**, *2015*, 1–11. [CrossRef]
73. Ali, N.; Khan, E.U.; Akhter, P.; Hussain, S. Wet depositional fluxes of ^{210}Pb- and ^{7}Be-bearing aerosols at two different altitude cities of North Pakistan. *Atmos. Environ.* **2011**, *45*, 5699–5709. [CrossRef]
74. Imran, A.; Zaman, Q.; Afzal, M. Temporal trends in the peak monsoonal precipitation events over Northeast Pakistan. *Pakistan J. Meteorol. Vol.* **2013**, *10*, 19–30.
75. Madhura, R.K.; Krishnan, R.; Revadekar, J.V.; Mujumdar, M.; Goswami, B.N. Changes in western disturbances over the Western Himalayas in a warming environment. *Clim. Dyn.* **2014**, *44*, 1157–1168. [CrossRef]
76. Rain Measurement. 2018. Available online: http://wiki.sandaysoft.com/a/Rain_measurement (accessed on 10 March 2018).
77. Ali, A.F.; Xiao, C.; Anjum, M.N.; Adnan, M.; Nawaz, Z.; Ijaz, M.W.; Sajid, M.; Farid, H.U. Evaluation and comparison of TRMM multi-satellite precipitation products with reference to rain gauge observations in Hunza River Basin, Karakoram Range, Northern Pakistan. *Sustainability* **2017**, *9*, 1954. [CrossRef]
78. 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]
79. Muhammad, W.; Yang, H.; Lei, H.; Muhammad, A.; Yang, D. Improving the Regional Applicability of Satellite Precipitation Products by Ensemble Algorithm. *Remote Sens.* **2018**, *10*, 577. [CrossRef]
80. 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]
81. Tokay, A.; Wolff, D.B.; Wolff, K.R.; Bashor, P. Rain gauge and disdrometer measurements during the Keys Area Microphysics Project (KAMP). *J. Atmos. Ocean. Technol.* **2003**, *20*, 1460–1477. [CrossRef]

remote sensing

MDPI

Article

Validation of the First Years of GPM Operation 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, Athens, GR 15236, Greece; katsanos@noa.gr
[2] Department of Meteorology, Nicosia, CY 1086, Cyprus; ftymvios@dom.moa.gov.cy
[3] The Cyprus Institute, Nicosia, CY 2121, Cyprus; s.michaelides@cyi.ac.cy
* Correspondence: adrianr@noa.gr; Tel.: +30-210-810-9201

Received: 27 July 2018; Accepted: 20 September 2018; Published: 21 September 2018

Abstract: Global Precipitation Measurement (GPM) high-resolution product is validated against rain gauges over the island of Cyprus for a three-year period, starting from April 2014. The precipitation estimates are available in both high temporal (half hourly) and spatial (10 km) resolution and combine data from all passive microwave instruments in the GPM constellation. The comparison performed is twofold: first the GPM data are compared with the precipitation measurements on a monthly basis and then the comparison focuses on extreme events, recorded throughout the first 3 years of GPM's operation. The validation is based on ground data from a dense and reliable network of rain gauges, also available in high temporal (hourly) resolution. The first results show very good correlation regarding monthly values; however, the correspondence of GPM in extreme precipitation varies from "no correlation" to "high correlation", depending on case. This study aims to verify the GPM rain estimates, since such a high-resolution dataset has numerous applications, including the assimilation in numerical weather prediction models and the study of flash floods with hydrological models.

Keywords: GPM; IMERG; TRMM; precipitation; Cyprus

1. Introduction

Several researchers have presented the comparison of satellite estimates with rain-gauge measurements in various study areas. One of the most common difficulties in such an attempt is the fact that satellite estimates refer to precipitation amounts over areas of a few kilometers around a gauge station, whereas rain-gauges measurements refer to the precipitation recorded in an area of a few centimeters in diameter (point source).

A literature overview of precipitation estimates from satellite, ground-based and gridded data could be found in [1–4]. However, recently released IMERG (Integrated Multi-satellite Retrievals) is going to reduce this sampling issue such that Global Precipitation Measurement (GPM) constellation satellites will have improved spatial and temporal resolutions (i.e., 30 min) [5,6].

Contrary to other satellites, such as its ancestor Tropical Rainfall Measuring Mission (TRMM), that could not measure light rain or snowfall, GPM-IMERG uses numerous sensors that measure at various frequencies from different satellites to detect both light and heavy rain and also snowfall. GPM expands TRMM's reach not only in terms of global coverage, but also through sophisticated satellite instrumentation, the inter-calibration of datasets from other microwave radiometers, coordinated merged precipitation datasets, reduced latency for delivering data products, simplified data access, expanded global ground-validation efforts, and integrated user applications [6]. GPM improves on TRMM's capabilities, although it has only two instruments versus the five instruments on TRMM. Dual-frequency Precipitation Radar (DPR) and the radiometer called GPM Microwave Imager (GMI) are some of the most advanced ever developed to measure precipitation from space. The DPR is the

only dual frequency radar in space and is capable of creating 3-D profiles and intensity estimates of precipitation ranging from rain to snow and ice. The GMI has a greater frequency range than TRMM's (13 channels versus 9 channels), which allows GPM to measure precipitation intensity and type through all cloud layers using a wider data swath.

Sharifi et al. [7] compared precipitation estimates of GPM IMERG final run data sets (along with TRMM 3B42 and ERA-Interim) and with meteorological synoptic stations' data over Iran to evaluate the GPM-based precipitation outputs for one-year period (March 2014 to February 2015). They found that IMERG underestimates precipitation slightly when considering days with observed precipitation. IMERG yields better results for detection of precipitation events on the basis of Probability of Detection (POD), Critical Success Index (CSI) and False Alarm Ratio (FAR) in areas with stratiform and orographic precipitation compared with other satellite/model data sets. In case of precipitation below 15 mm/day IMERG indicated better results, while for heavy precipitation (>15 mm/day) IMERG was far superior to the other products.

Manz et al. [8] performed validation of the GPM IMERG Day-1 and TRMM TMPA products against rain gauge stations across Ecuador and Peru from March 2014 to August 2015 at different spatial (0.1°, 0.25°) and temporal (1 h, 3 h, daily) scales. They found that IMERG showed better detection and quantitative precipitation intensity estimation ability than TMPA, particularly in the high Andes. Furthermore, IMERG showed better characterization of gauge observations after separating precipitation detection and precipitation rate estimation.

Prakash et al. [9] presented an assessment of IMERG, TMPA and Global Satellite Mapping of Precipitation (GSMaP, v.6) estimates against gauge-based observations over India at a daily scale for the southwest monsoon season (June to September 2014). Results indicated that the IMERG product represented well the mean monsoon precipitation and its variability, performing a clear improvement in missed and false precipitation bias over India. However, in areas that precipitation is dominated by orographic effects (northeast India) and over rain-shadow regions (southeast peninsular India) all three satellite-based precipitation estimates did not perform well (in terms of correlation coefficient, RMSE, bias).

Sungmin et al. [10] presented an analysis of the IMERG (v.3) Early, Late, and Final (IMERG-E, IMERG-L, and IMERG-F) half-hourly precipitation estimates against gauge-based gridded precipitation data from a high-density climate station network in southeastern Austria for the period of April–October in 2014 and 2015. Results showed that IMERG-F precipitation estimates were in the best overall agreement with the gauge-based gridded precipitation data, followed by IMERG-L and IMERG-E estimates, particularly for the hot season.

Tan and Duan [11] evaluated GPM IMERG and TRMM Multisatellite Precipitation Analysis (TMPA) 3B42 and 3B42RT precipitation data over Singapore on daily, monthly, seasonal and annual scales from April 2014 to January 2016. They found that both IMERG and TMPA products were well correlated with gauges data at monthly scale, but moderately at daily scale (IMERG had the lowest systematic bias). However, IMERG had better performance in the characterization of spatial precipitation variability and precipitation detection capability compared to the TMPA products. All satellite precipitation products overestimated moderate precipitation events (1–20 mm/day) and underestimated light (0.1–1 mm/day) and heavy (>20 mm/day) precipitation events. They concluded that in overall, IMERG showed a slight improvement compared to the TMPA products over Singapore.

Asong et al. [12] evaluated GPM IMERG (V03 Final Run) data against ground-based measurements at the 6-hourly, daily, and monthly time scales over different terrestrial ecozones of southern Canada during March 2014 to January 2016. They found that considering the mean daily precipitation IMERG and ground-based records showed similar regional variations, while IMERG was overestimated during higher monthly precipitation amounts. Regarding daily and 6-hourly time scales IMERG and ground-based records found to be in agreement for the months of June–September, unlike months November–March. Considering precipitation extremes, it was found that generally

IMERG captured well the distribution of heavy precipitation amounts, although correlation was spatially varied.

Tang et al. [13] evaluated GPM IMERG Day-1 product over Mainland China from April to December 2014 (hourly data) against TRMM TMPA (3B42V7) product (3-hourly and daily data) data and hourly ground-based observations. They found that Day-1 IMERG showed better performance than 3B42V7 at both sub-daily and daily timescales, with significant differences between the two products noticed at the sub-daily resolution; IMERG especially performed better at the mid- and high-latitudes and at relatively dry climate regions; Regarding ability of capturing precipitation intensity, IMERG reproduced better the probability density function (PDF), especially in the low ranges.

Tian et al. [14] investigates the dependency of the evaluation of the GPM IMERG precipitation product on the gauge density of a ground-based rain gauge network as well as precipitation intensity over five sub-regions in mainland China during two warm seasons (April–October 2014 and 2015). They found that the denser the rain gauge network the better the evaluation metrics. In addition, IMERG tends to overestimate (underestimate) light (heavy) precipitation events.

Xu et al. [15] evaluated two satellite precipitation products (GPM IMERG and TRMM 3B42V7) in southern Tibetan Plateau region focusing on the effects of topography and precipitation intensity, for the rainy season from May to October 2014. Results indicated that GPM product presented better ability in detecting daily precipitation accumulation at all spatial scales and elevation ranges. It also tends to underestimate the amount of light rain (0–1 mm/d) events, while showing better detecting ability for light precipitation (0–5 mm/d) events; however, the detecting ability at high-elevation (>4500 m) regions was not defined. In terms of the ability of precipitation detection, false alarm ratio (FAR) of GPM (14%) was lower than that of TRMM (21%), while missing ratio of GPM (13%) was larger than that of TRMM (9%).

Dezfuli et al. [16] evaluated GPM IMERG (V04A) precipitation data using in situ observations from selected stations in West and East Africa, along with some comparisons between IMERG and TMPA. Results indicated temporal (by season) and spatial (by region) variation of correlation between the satellite-based products and stations data. They also found that precipitation diurnal cycle is relatively better captured by IMERG than TMPA, while both products demonstrated better agreement with stations data in East Africa and humid West Africa than in the southern Sahel. Regarding annual cycle, differences were found during the short rains in East Africa. A direct comparison between IMERG and TMPA over the entire continent was also established concluding that IMERG and TMPA presented large differences over mountainous regions.

Liu [17] commented on the comparison between the IMERG Final Run monthly product and the TMPA monthly product (3B43) in the boreal summer of 2014 and the boreal winter of 2014/15 on a global scale. It was found that IMERG monthly product could capture major heavy precipitation regions in the Northern and Southern Hemispheres reasonably well. However, differences between IMERG and 3B43 varied with surface types and precipitation rates in both seasons, with systematic differences found much smaller over land. When low precipitation rates were considered positive relative differences (IMERG > 3B43) were noticed in contrast to high precipitation rates (IMERG < 3B43). Over ocean, under all precipitation rates negative differences were found. Tan and Santo [18] studied the performance of six satellite precipitation products (SPPs), GPM IMERG products (IMERG-E, IMERG-L and IMERG-F), TMPA 3B42 and 3B42RT products and the PERSIANN-CDR product, against precipitation gauges for an almost two years period (March 2014 to February 2016), over Malaysia. Results indicated that all the SPPs performed well in annual and monthly precipitation measurements. However, moderate correlation was found at daily precipitation estimations.

Regarding precipitation detection ability, all the SPPs underestimated light (0–1 mm/day) and extreme (>50 mm/day) precipitation events, while overestimated moderate and heavy (1–50 mm/day) precipitation events. The IMERG demonstrated better ability in detecting light precipitation (0–1 mm/day) compared to the other SPPs.

In the literature, only a few studies with evaluation of satellite precipitation products over Cyprus are reported based on gauge observations and a few satellite precipitation products.

Retalis et al. [19] 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 for a 15-year period (1998–2012). They found that TRMM and E-OBS estimates were well fitted with the gauge data, although they were both influenced by the variation of elevation.

A study for a 30-year period (1981–2010) for the precipitation database Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) was performed over the island of Cyprus [20]. CHIRPS database was evaluated against the dense and reliable network of rain gauges. Results showed good correlation between CHIRPS values and recorded precipitation; although an overestimation, in the CHIRPS values, of the in-situ precipitation data has been noted during the last decade.

Katsanos et al. [21] in their study calculated a number of climatic indices for the CHIRPS precipitation dataset and records from rain gauges on high spatial (0.05°) and temporal (daily) resolution for a period of 30 years (1981–2010). Results indicates a quite 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.

Retalis et al. [22] proposed a novel downscaling method of TRMM 3B42 product in order to attain high resolution (1 km $\times$ 1 km) precipitation datasets, by correlating the CHIRPS dataset with elevation information and the normalized difference vegetation index (NDVI) from satellite images at 1 km $\times$ 1 km, utilizing artificial neural network models. The resulted product was validated with precipitation measurements from the corresponding rain gauge network of the Cyprus Department of Meteorology. It was found that the new product was able to capture the spatial precipitation variability, especially in the mountainous areas. Its overall accuracy in the annual products was considered rather significant, while the monthly product, although characterized by variability in the range of correlation (r) values, indicated rather significant values when mean monthly products were considered.

Katsanos et al. [23] analyzed the daily precipitation (rain gauges) database for the island of Cyprus for period of 50 years. They calculated a number of climatic indices for precipitation, including the Standardized Precipitation Index (SPI for the recent 30-year period 1981–2010 and compared the results with those of the period 1961–1990. The results showed an increase in the occurrence of heavy precipitation events and a slight decrease in extreme drought events.

The aim of this study is to assess the accuracy of the new GPM IMERG satellite precipitation product over Cyprus on a monthly basis and during extreme events. To our knowledge, no such report exists yet to study the GPM constellation satellites data over Cyprus.

2. Data and Methodology

2.1. Study Area

The island of Cyprus is located in the north-eastern corner of the Mediterranean Sea. It is characterized by a typical eastern-Mediterranean climate, with wet winters and dry summers, in terms of precipitation. With an average elevation of 91 m above Mean Sea Level, 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 narrow Pentadaktylos range, rising to 900 m and bordering the northern coast. Between the two mountain ranges, lies the central plain and along their seaward margins a more or less narrow coastal strip (Figure 1). The highest average annual precipitation values are recorded at the meteorological stations located on the mountainous areas of the island and the lowest at the coastal ones.

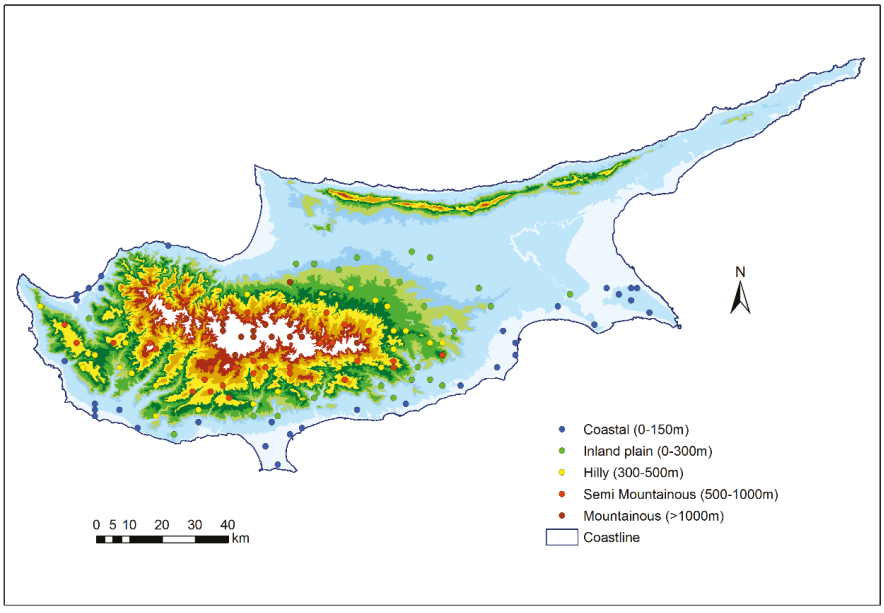


Figure 1. The island of Cyprus, with the location of the stations used; stations are grouped in terms of their elevation.

2.2. Data

2.2.1. In-Situ Rain Gauge Data

The in-situ measurements of meteorological stations operated by the Cyprus Department of Meteorology were applied for ground validation. The network is considered reliable with respect to the quality of the measuring stations and routine maintenance, dense having on the average 1 station per 30 km^2, it is well-distributed covering all different climatic zones of the island (Figure 1) and provides historical daily rainfall measurements for more than a century (Figure 2). It is currently comprised of three synoptic stations, 35 online stations, 20 automatic weather stations, 90 ordinary rain gauges and 25 analog rain recorders whilst an ongoing effort exists concerning the replacement of the analog instruments with digital. All measurements are subject to quality control by qualified personnel. Daily and hourly data records from 136 rain gauges with complete data series of daily values during the study period were selected.

The area average precipitation provided in Figure 2 illustrates the climatology of the annual precipitation for the entire time series of the recorded data and it clearly demonstrates a decreasing trend in the annual average precipitation over the climatic periods of 1941–1970, 1951–1980, 1961–1990 and 1971–2000. The annual average area precipitation was calculated from monthly precipitation measurements with planimeter while the method was replaced by GIS after 2006. A two-year period of evaluation of the two methods revealed no significant deviations amongst them, differentiating below 4% of the yearly calculated value. This method of area calculation of precipitation is not considered accurate for the entire area since it does not take under consideration the altimeter difference of the measurement stations, but it is a good indicator for the yearly amount of precipitation received over the entire island. The total annual precipitation is approximately 300–400 mm year^{-1} [24]. The maximum area average precipitation is 800 mm, recorded in 1968/1969; the minimum area average precipitation is 213 mm, recorded in 1972–1973.

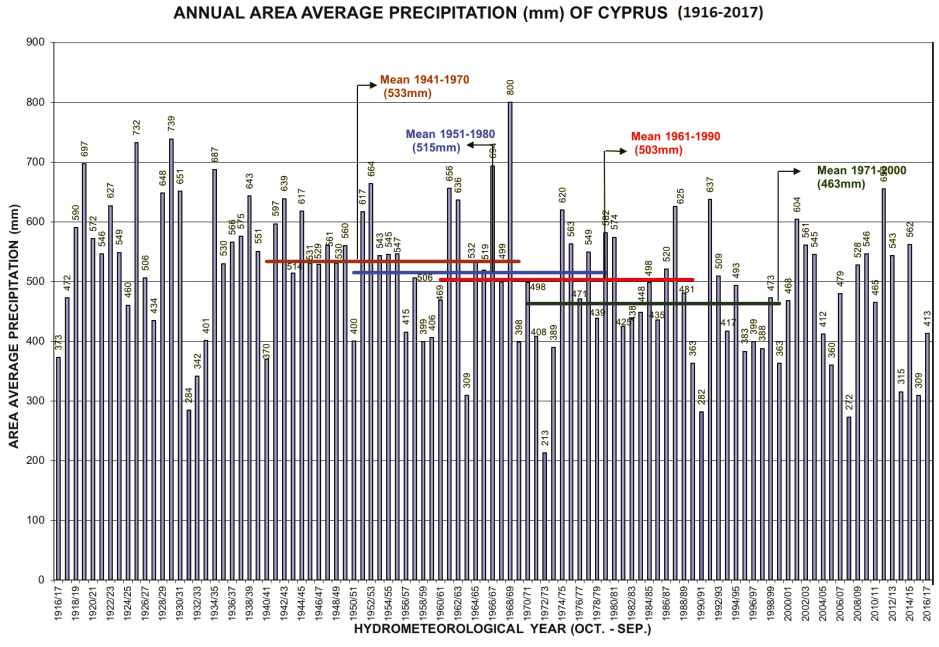


Figure 2. The annual average precipitation of Cyprus for the period 1916–2017.

2.2.2. IMERG Data

Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG) products provide quasi-global (60° N–60° S) precipitation estimates, beginning March 2014, from the combined use of passive microwave (PMW) and infrared (IR) satellites comprising the GPM constellation. The IMERG Final data used in this study (version 04A) have a temporal resolution of 1/2 h and a spatial resolution of 0.1 degree. The IMERG-F product ingests GPCC monthly gauge analyses making it a very reliable dataset. The period of study spans from April 2014 to February 2017. A known problem exists regarding a low bias in GPROF algorithm at high latitudes due to the low sensitivity of the DPR to light rain and snow, which was not fully fixed in the latest version (e.g., 05), despite the modification of giving more weight to GMI data. Furthermore, in version V05 IMERG Final datasets have fewer amounts of microwave-based data, compared to versions 03 and 04, since Version 05 GPROF-TMI estimates had not been computed for the GPM era. Another known issue is the resolution difference between GPCC (1 deg) and IMERG (0.1 deg) that lead to an unphysical "blockiness" after the bilinear interpolation used at the first versions, that was corrected at V05, after the removal of the GPCC grid box volume adjustment.

It should be noted that for each day GPM provides 48 files with measurements which means that, for a 3-year period, one has to handle (download and process) more than 50,000 files.

The satellite products were validated against the gauge stations data at daily and monthly time scales. Precipitation data within each satellite cell was compared directly. In cases where precipitation data of more than one gauges was included within the satellite cell, the comparison was carried out based on the average of the gauge stations data.

The measured (gauge stations) and estimated (IMERG) precipitation pairs were also used to construct a 2 × 2 contingency table (Table 1) in order to investigate how well the IMERG estimates correspond to the measured gauge data.

Table 1. Contingency table used for calculating the statistical scores.

		Gauges		
		Yes	No	Total
GPM IMERG	Yes	Hits	False alarms	Forecast Yes
	No	Misses	Correct negatives	Forecast No
	Total	Observed Yes	Observed No	Total

2.3. Statistical Scores

Next, some statistical scores were calculated [25].

Bias score (frequency bias)

$$\text{BIAS} = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}} \tag{1}$$

BIAS ranges from 0 to ∞, with perfect score to be equal to 1. BIAS measures the ratio of the frequency of forecast events to the frequency of observed events. It indicates whether the forecast system has a tendency to under-forecast (BIAS < 1) or over-forecast (BIAS > 1) events. It does not measure how well the forecast corresponds to the observations, it only measures relative frequencies.

Probability of detection (hit rate, POD)

$$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} \tag{2}$$

POD ranges from 0 to 1, with perfect score to be equal to 1. POD is sensitive to hits but ignores false alarms. It is very sensitive to the climatological frequency of the event. It is also a good measure for rare events. It can be artificially improved by issuing more "yes" forecasts to increase the number of hits. It is better to be used in conjunction with the false alarm ratio (below). False alarm ratio (FAR)

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \tag{3}$$

FAR ranges from 0 to 1, with perfect score to be equal to 0. FAR is sensitive to false alarms but ignores misses. It is very sensitive to the climatological frequency of the event. It is suggested to be used in conjunction with the probability of detection (above).

Probability of false detection (POFD)

$$\text{POFD} = \frac{\text{false alarms}}{\text{correct negatives} + \text{false alarms}} \tag{4}$$

POFD ranges from 0 to 1, with perfect score equal to 0. POFD is sensitive to false alarms but ignores misses. It can be artificially improved by issuing fewer "yes" forecasts to reduce the number of false alarms.

Success ratio (SR)

$$\text{SR} = \frac{\text{hits}}{\text{hits} + \text{false alarms}} \tag{5}$$

SR ranges from 0 to 1, with perfect score equal to 1. SR gives information about the likelihood of an observed event, given that it was forecast. It is sensitive to false alarms but ignores misses. SR is equal to 1-FAR. POD is plotted against SR in the categorical performance diagram.

Threat score (critical success index, TS)

$$\mathrm{TS} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}} \tag{6}$$

TS range from 0 to 1, with 0 indicating no skill and perfect score equal to 1. TS measures the fraction of observed and/or forecast events that were correctly predicted. It can be thought of as the accuracy when correct negatives have been removed from consideration, that is, TS is only concerned with forecasts that count. Hanssen and Kuipers discriminant score (HK)

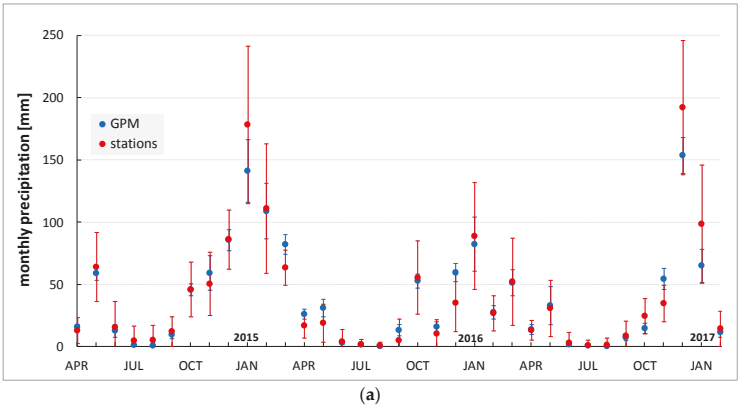
$$\mathrm{HK} = \frac{\text{hits}}{\text{hits} + \text{misses}} - \frac{\text{false alarms}}{\text{false alarms} + \text{correct negatives}} \tag{7}$$

HS ranges from -1 to 1, with 0 indicating no skill and perfect score equal to 1. HS uses all elements in contingency table. It does not depend on climatological event frequency. The expression is identical to HK = POD − POFD, but the Hanssen and Kuipers score can also be interpreted as (accuracy for events) + (accuracy for non-events) − 1.

3. Results and Discussion

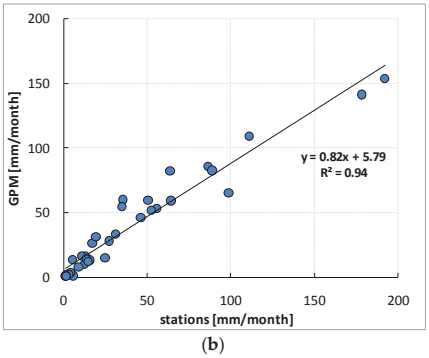
3.1. Monthly Validation

At first, the comparison of the monthly precipitation gauge data and IMERG data for the study period was performed. The monthly variation of the two data sets along with the corresponding standard variation are shown in Figure 3a. It is noticed that both data have a similar pattern of monthly variation, however, during the rainy period (October to February), IMERG was overestimating precipitation. The scatter plot of the respective data and the calculated coefficient of determination (r^2 = 0.93) are presented in Figure 3b. In order to further explore the effect of topography on this relationship, data were categorized in five classes according to the elevation of gauge stations. A color code is adopted in Figure 1 in order to display the five different classes used in the present study. The five classes are shown in Table 2, together with the number of stations falling in each class. The average elevation of all the stations that were used is 368 m.



(a)

Figure 3. *Cont.*



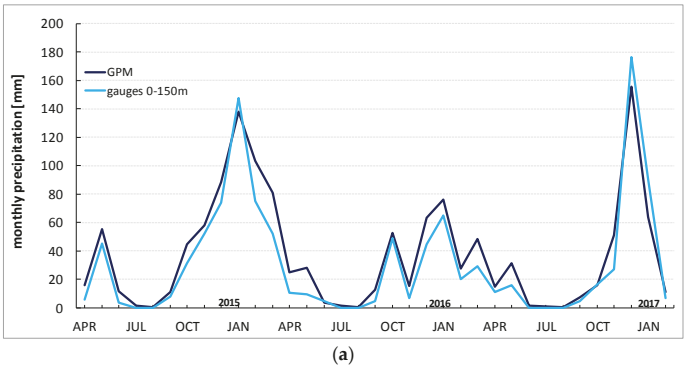
(**b**)

Figure 3. Monthly precipitation from GPM (average of 84 grid cells) versus gauges (average of 136 stations) from April 2014 to February 2017, for all stations (**a**), and the corresponding scatter plot (**b**). Upper and lower limits of the monthly precipitation bars refer to corresponding standard deviation.

Table 2. The five classes of rain gauges. Categorization is made on the basis of the rain gauge's elevation (see Figure 1). The number of rain gauges along with the corresponding GPM grid cells in each class are also shown.

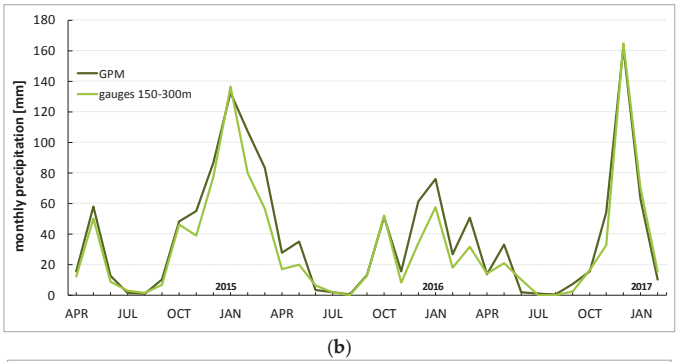
Class No.	Elevation (m)	Categorization	Number of Rain Gauges	Number of GPM Grid Cells
1	0–150	Coastal	42	26
2	150–300	Inland plain	18	14
3	300–500	Hilly	25	17
4	500–1000	Semi Mountainous	37	20
5	>1000	Mountainous	14	7

It is noticed that both IMERG and gauge data depict similar patterns in terms of monthly variation in all cases, regardless of the elevation of the stations. For the cases that elevation lower than 500 m is considered, small differences are observed, while the corresponding coefficients of determination are within the range of 0.92–0.94 (Figures 4 and 5). However, when elevation higher than 500 m is considered GPM tends to underestimate precipitation with corresponding coefficients of determination of 0.92 (500–1000 m) and 0.87 for elevation higher than 1000 m. This is more evident during the wet period (December, January), where differences noticed are larger, especially in the case where elevation is higher than 1000 m.

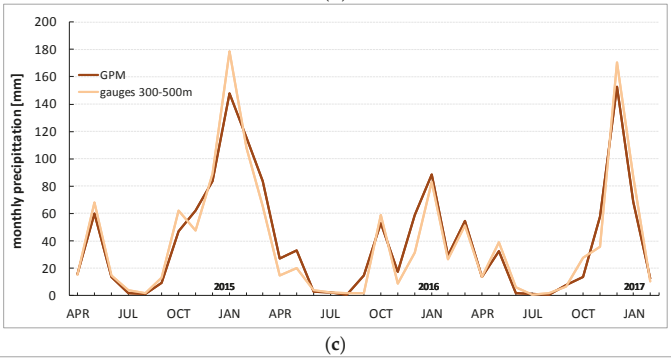


(**a**)

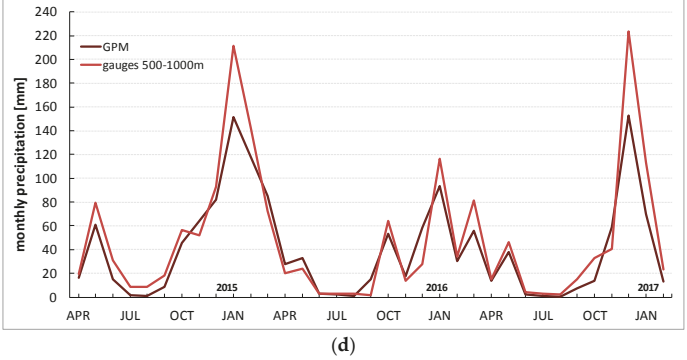
Figure 4. *Cont.*



(**b**)

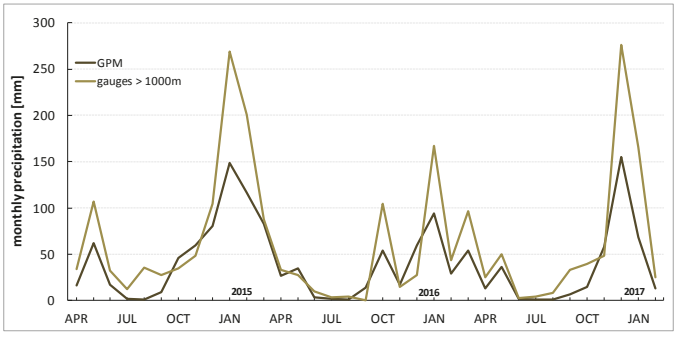


(**c**)



(**d**)

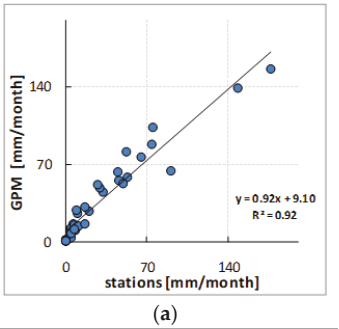
Figure 4. *Cont.*



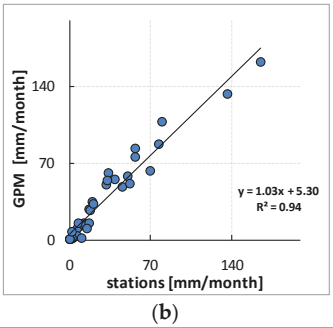
(**e**)

Figure 4. Monthly precipitation from GPM versus gauges from April 2014–February 2017, for stations with elevation from (**a**) 0 to 150 m; (**b**) 150 to 300 m; (**c**) 300 to 500 m; (**d**) 500 to 1000 m; (**e**) greater than 1000 m.

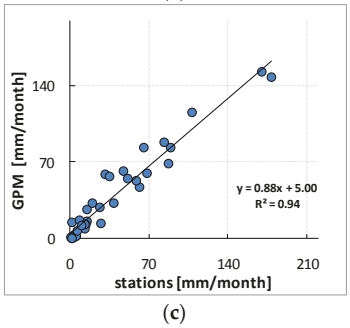
Accordingly, the root mean squared error (RMSE) and mean absolute error (MAE) metrics were also calculated (Table 3). It is noticed that RMSE ranges from 11.82 to 48.75 and MAE from 8.62 to 30.43 when is calculated for each cell and for the whole period of study. The maximum value refers to a cell within which reside stations with elevation greater than 1100 m, while the second maximum is also in a similar cell regarding elevation. On the other hand, the minimum values of both metrics refer to a cell where a station of almost 600 m is attached.



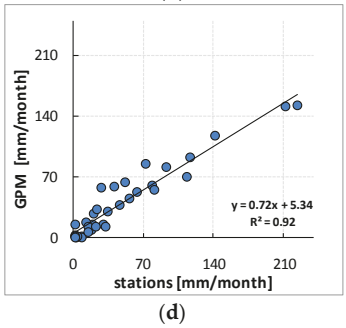
(**a**)



(**b**)

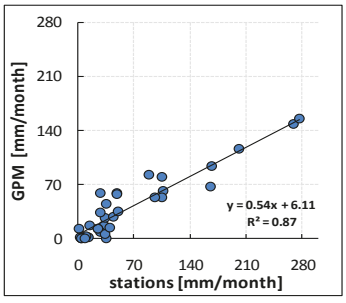


(**c**)



(**d**)

Figure 5. *Cont.*



(**e**)

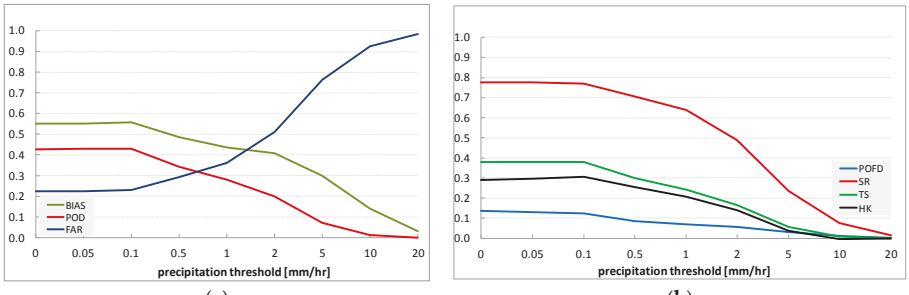
Figure 5. Scatter plots of monthly precipitation from GPM versus gauges from April 2014–February 2017, for stations with elevation from (**a**) 0 to 150 m; (**b**) 150 to 300 m; (**c**) 300 to 500 m; (**d**) 500 to 1000 m; (**e**) greater than 1000 m.

Table 3. Results of RMSE and MAE.

	RMSE (Cell)	MAE (Cell)	RMSE (Month)	MAE (Month)
Average	23.80	15.65	19.86	15.65
Minimum	11.82	8.62	2.47	1.31
Maximum	48.75	30.43	69.03	59.78

The month with the maximum values of RMSE and MAE is December 2016 with 69.03 and 59.78 respectively, while as expected, all winter months which correspond to the wet period of the hydrological year, have high values of both metrics. On the other hand, August 2015 has the lowest values, 2.47 and 1.31 respectively, and the summer months in general have low values, as they correspond to the dry period of the year (May–August).

The statistical scores based on contingency tables were calculated comparing hourly values of gauge records and IMERG estimates during the study period considering that the daily (24 h) precipitation values were greater than 10 mm. Results are presented in Figure 6.



(**a**) (**b**)

Figure 6. Statistical scores based on contingency tables: BIAS, POD, FAR (**a**); POFD, SR, TS, HK (**b**).

Correct negatives increase significantly as precipitation threshold increases; misses and hits are drastically decreased with increasing precipitation threshold, while false alarms are also reduced. It is noticed that for precipitation threshold within the range 0 to 0.1 mm/h, FAR is relatively low (~0.2), while POD and BIAS are at their maximum (0.4 and 0.55, respectively). Accordingly, as precipitation

threshold increases, FAR also increases, while POD and BIAS decrease. On the other hand, POFD, SR, TS and HK decrease with increasing precipitation threshold.

Figure 7a illustrates the frequency distribution (bars) of the correlation coefficient values for all cases (24 h precipitation > 10 mm) recorded during the first years of GPM over Cyprus together with the cumulative frequency (line) for the same values. The pairs that are compared are the GPM grid-cells with precipitation estimates and the corresponding gauges values, both as hourly rates during the whole duration of the events, since hourly is the highest time resolution that the observations can be available.

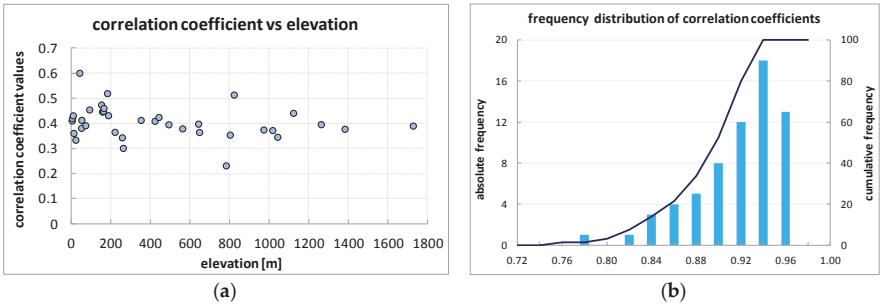


Figure 7. Frequencies of correlation coefficients (**a**) and frequency distribution of correlation coefficients (**b**).

It is obvious that there is not any value with a clear dominance since almost all coefficient values lay between 3% and 5%, however there are very few cases with negative correlation. This can be seen also at the cumulative line, where about 10% of the values are negative, 60% are greater than 0.30 and almost 30% are greater than 0.60. High values i.e., 0.75 or greater, have about 20% of the studied cases.

Figure 7b illustrates the frequency distribution of correlation coefficient values. Breaking down to 65 grid cells, the correlation coefficient values between GPM estimates and the corresponding monthly observed precipitation (by one or more gauges) vary between 0.78 and 0.96. However, the vast majority of the values is over 0.90.

3.2. Daily Validation

In order to study the ability of IMERG data to detect the daily variation of precipitation, we analyzed the whole dataset aiming mainly to extreme precipitation events recorded during the period of study. Accordingly, we proceed to the selection of a few case study events during the three years period. Figure 8 presents the time propagation of the precipitation measured by the stations (average value for all stations that recorded precipitation) against that estimated by the GPM (again an average value for all grid cells), for the entire duration of each event. The GPM values are converted from half-hourly to hourly, in order to match the time resolution of the station data.

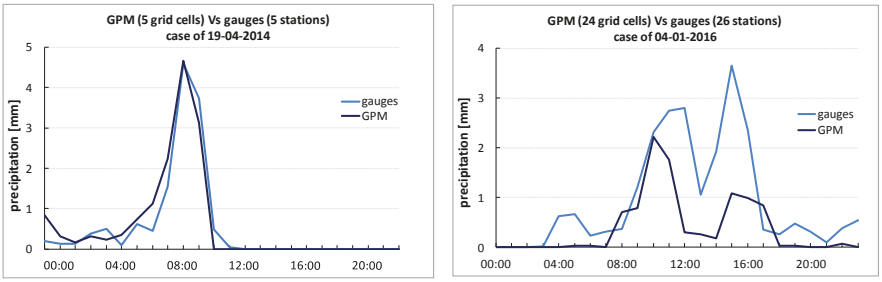


Figure 8. *Cont.*

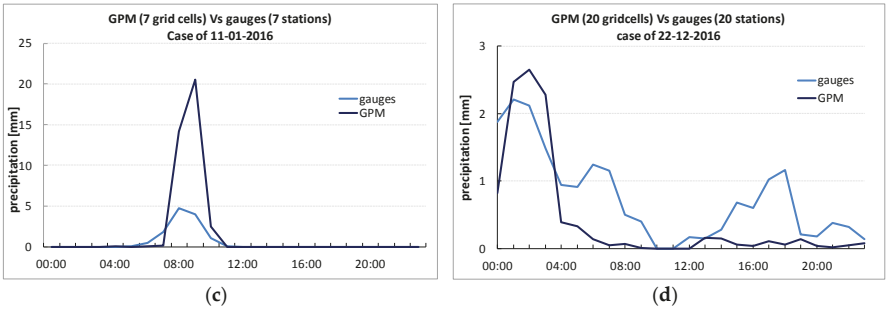


Figure 8. Time propagation of the precipitation recorded by stations (light blue) and that estimated by GPM (dark blue), for four selected cases dated 19 April 2014 (**a**); 04 January 2016 (**b**); 11 January 2016 (**c**); 22 December 2016 (**d**).

The examples shown in Figure 8, illustrate that there is a case dependency in the ability of GPM to estimate precipitation on high temporal resolution. Indeed, there are cases where the IMERG estimates match excellent the recorded precipitation (Figure 8a) or follow well the time evolution but miss the height of precipitation (Figure 8b), but in most of the cases precipitation is estimated just satisfactory (Figure 8c,d).

4. Conclusions

This study, as the title indicates clearly, is a first attempt to validate a specific GPM precipitation product by using in situ surface observations over the island of Cyprus. In comparing the two datasets, neither an upscaling of the rain-gauge measurements to the GPM grid nor an interpolation of the GPM measurements to the station level was attempted. The satellite product was validated against the gauge data by simply comparing the former with the latter within each satellite cell. The limitation of the data span is not encouraging more rigorous validation approaches. Upscaling the rain-gauge measurements or interpolating the satellite product to the station sites could be an alternative as more data become available but in such a case associated issues should also be considered [26].

An evaluation of the IMERG product in daily and monthly scales against rain-gauges from the dense and reliable network operated by the Cyprus Department of Meteorology in Cyprus from April 2014 to February 2017 was performed. Based on the statistical analysis presented, it was concluded that, in overall, a very good agreement on monthly level was established, although IMERG tend to be underestimated as elevation goes higher, especially for areas higher than 1000 m elevation. Regarding

daily evaluation, in terms of extreme recorded events, it seems that there exists a case dependency, while no specific correlation with elevation was established.

In summary, that accurate estimation by satellite-based precipitation products remains still a challenge. However, form this preliminary assessment on the IMERG product, it is highlighted that it could be useful for use in hydrological studies of small basins and flood studies due to its finer spatial and temporal resolution than TMPA.

The research in the field is still open. The authors aim to further evaluate the product, while implementing the new released version of TMPA product to the spatial resolution of IMERG.

Author Contributions: All authors contributed extensively to the work presented in this paper. The manuscript was prepared by A.R. and revised by D.K., F.T. and S.M., D.K. and F.T. provided and analyzed data. All authors discussed the results and implications of the manuscript at all stages.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge the provision of the rain gauge data by the Cyprus Department of Meteorology and the GPM data by the National Aeronautics and Space Administration, U.S.A.

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

References

1. 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.; de Castro, M. Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.* **2012**, *104–105*, 70–97. [CrossRef]
2. Tapiador, F.J.; Navarro, A.; Levizzani, V.; García-Ortega, E.; Huffman, G.J.; Kidd, C.; Kucera, P.A.; Kummerow, C.D.; Masunaga, H.; Petersen, W.A.; et al. Global precipitation measurements for validating climate models. *Atmos. Res.* **2017**, *197*, 1–20. [CrossRef]
3. 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]
4. Michaelides, S.; Tymvios, F.S.; Michaelidou, T. Spatial and temporal characteristics of the annual rainfall frequency distribution in Cyprus. *Atmos. Res.* **2009**, *94*, 606–615. [CrossRef]
5. 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]
6. 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]
7. 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]
8. Manz, B.; Páez-Bimos, S.; Horna, N.; Buytaert, W.; Ochoa-Tocachi, B.; Lavado-Casimiro, W.; Willems, B. Comparative Ground Validation of IMERG and TMPA at Variable Spatio temporal Scales in the Tropical Andes. *J. Hydrometeor.* **2017**, *18*, 2469–2489. [CrossRef]
9. Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Liu, Z.; Norouzi, H.; Pai, D.S. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.* **2018**, *556*, 865–876. [CrossRef]
10. 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]
11. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM Precipitation Products over Singapore. *Remote Sens.* **2017**, *9*, 720. [CrossRef]
12. Asong, Z.E.; Razavi, S.; Wheater, H.S.; Wong, J.S. Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over Southern Canada against ground Precipitation Observations: A Preliminary Assessment. *J. Hydrometeor.* **2017**, *18*, 1033–1050. [CrossRef]

13. 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]
14. Tian, F.; Hou, S.; Yang, L.; Hu, H.; Hou, A. How Does the Evaluation of the GPM IMERG Rainfall Product Depend on Gauge Density and Rainfall Intensity? *J. Hydrometeor.* **2018**, *19*, 339–349. [CrossRef]
15. 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]
16. 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]
17. Liu, Z. Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results. *J. Hydrometeor.* **2016**, *17*, 777–790. [CrossRef]
18. 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]
19. Retalis, A.; Katsanos, D.; Michaelides, S. Precipitation climatology over the Mediterranean Basin—Validation over Cyprus. *Atmos. Res.* **2016**, *169*, 449–458. [CrossRef]
20. 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]
21. 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]
22. Retalis, A.; Tymvios, F.; Katsanos, D.; Michaelides, S. Downscaling CHIRPS precipitation data: An artificial neural networks modelling approach. *Int. J. Remote Sens.* **2017**, *38*, 3943–3959. [CrossRef]
23. Katsanos, D.; Retalis, A.; Tymvios, F.; Michaelides, S. Study of extreme wet and dry periods in Cyprus using climatic indices. *Atmos. Res.* **2018**, *208*, 88–93. [CrossRef]
24. 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]
25. Jolliffe, I.T.; Stephenson, D.B. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*; John Wiley & Sons: Chichester, UK, 2003; p. 240.
26. Tyralis, H.; Koutsoyiannis, D. On the prediction of persistent processes using the output of deterministic models. *Hydrol. Sci. J.* **2017**, *62*, 2083–2102. [CrossRef]

remote sensing

MDPI

Article

Analysis of Livorno Heavy Rainfall Event: Examples of Satellite-Based Observation Techniques in Support of Numerical Weather Prediction

Elisabetta Ricciardelli [1,*], Francesco Di Paola [1], Sabrina Gentile [1,2], Angela Cersosimo [1], Domenico Cimini [1,2], Donatello Gallucci [1], Edoardo Geraldi [1,3], Salvatore Larosa [1], Saverio Teodosio Nilo [1], Ermann Ripepi [1], Filomena Romano [1] and Mariassunta Viggiano [1]

[1] Institute of Methodologies for Environmental Analysis, National Research Council (IMAA/CNR), 85100 Potenza, Italy; francesco.dipaola@imaa.cnr.it (F.D.P.); sabrina.gentile@imaa.cnr.it (S.G.); angela.cersosimo@imaa.cnr.it (A.C.); domenico.cimini@imaa.cnr.it (D.C.); donatello.gallucci@imaa.cnr.it (D.G.); edoardo.geraldi@imaa.cnr.it (E.G.); salvatore.larosa@imaa.cnr.it (S.L.); saverio.nilo@imaa.cnr.it (S.T.N.); ermann.ripepi@imaa.cnr.it (E.R.); filomena.romano@imaa.cnr.it (F.R.); mariassunta.viggiano@imaa.cnr.it (M.V.)

[2] Center of Excellence Tele Sensing of Environment and Model Prediction of Severe events (CETEMPS), University of L'Aquila, 67100 L'Aquila, Italy

[3] Institute of Archeological Heritage, Monuments and Sites, National Research Council (IBAM/CNR), 85100 Potenza, Italy

* Correspondence: elisabetta.ricciardelli@imaa.cnr.it; Tel.: +39-097-142-7292

Received: 8 August 2018; Accepted: 24 September 2018; Published: 26 September 2018

Abstract: This study investigates the value of satellite-based observational algorithms in supporting numerical weather prediction (NWP) for improving the alert and monitoring of extreme rainfall events. To this aim, the analysis of the very intense precipitation that affected the city of Livorno on 9 and 10 September 2017 is performed by applying three remote sensing techniques based on satellite observations at infrared/visible and microwave frequencies and by using maps of accumulated rainfall from the weather research and forecasting (WRF) model. The satellite-based observational algorithms are the precipitation evolving technique (PET), the rain class evaluation from infrared and visible observations (RainCEIV) technique and the cloud classification mask coupling of statistical and physics methods (C-MACSP). Moreover, the rain rates estimated by the Italian Weather Radar Network are also considered to get a quantitative evaluation of RainCEIV and PET performance. The statistical assessment shows good skills for both the algorithms (for PET: bias = 1.03, POD = 0.76, FAR = 0.26; for RainCEIV: bias = 1.33, POD = 0.77, FAR = 0.41). In addition, a qualitative comparison among the three technique outputs, rain rate radar maps, and WRF accumulated rainfall maps is also carried out in order to highlight the advantages of the different techniques in providing real-time monitoring, as well as quantitative characterization of rainy areas, especially when rain rate measurements from Weather Radar Network and/or from rain gauges are not available.

Keywords: heavy precipitation; rainfall retrieval techniques; forecast model

1. Introduction

Extreme rainfall events represent a looming and great threat to human safety and economy. In fact, very heavy precipitation can cause loss of life and serious harm, including the destruction of buildings and infrastructure, such as roads, water systems, and power lines, negatively affecting the economy. In particular, floods are the predominant cause of weather-related disruption to the transport sector [1], while strong winds and lightning thunderstorms represent a high financial risk for the aviation industry [2,3]. The correct prediction of these extreme events and their continuous

monitoring are crucial for saving human lives, and to limit damage and economy loss. Numerical weather prediction (NWP) models are not always able to precisely identify magnitude, location, and time of the extreme rainfall events, because of their high variability and discontinuity in time and space [4]. Several algorithms, based on remote sensing observations developed in last years, may be able to support NWP models in detecting and characterizing spatially and temporally extreme rainfall events. In particular, in [5], a detection scheme of convective initiation is proposed by combining two existing algorithms [6,7] that use infrared observations from geostationary data for detecting and tracking convective events. Moreover, in [8], three machine learning approaches for convective initiation detection have been proposed. In [9], hyperspectral infrared observations from geostationary orbit are used to implement a framework for the prediction of local severe storms. The better accuracy of hyperspectral infrared observations in deep convection identification is also demonstrated in [10], where the difference in brightness temperature is investigated to assess the identification of uncertainty for different instruments. To this end, a case study related to airplane accident caused by a not correctly forecast deep convection event is analyzed. In [11], infrared and water vapor wavelengths geostationary observations are explored together with tropopause temperature information provided by an NWP model, to implement a method that aims to stratify a cloud shield to better characterize and track the evolution of convective clouds. Other works [12,13] propose the CMORPH (Climate Prediction Center morphing) method, that uses motion vectors derived from geostationary satellite infrared data to propagate forward in space and time the precipitation measurements retrieved by passive microwave (PMW) radiometers, in order to obtain global precipitation estimates at 8 km of spatial resolution every half hour. Moreover, ref. [14] demonstrated the reliability of different physics-based passive microwave (PMW) precipitation retrieval algorithms of the GPM (global precipitation measurement) mission in providing quantitative description of the precipitation for heavy rainfall events that occurred during the autumn 2014 in Italy. This paper intends to investigate the remote sensing monitoring value in supporting NWP forecast in characterizing the rainfall event evolution, when NWP fails to detect these extreme events as localized both in space and time. To this aim, three satellite based techniques are applied to the extreme rainfall event that hit Livorno on 9 and 10 September 2017, when a series of storms occurred in central Italy and very intense precipitation—more than 150 mm in few hours—was recorded by rain gauges placed in several locations in Tuscany and Lazio. In Livorno, the storm dumped more than twice the mean monthly rainfall in just 2 h (in detail, the rain gauge located in Livorno recorded about 60 mm of rainfall between 20:45 UTC and 22:00 UTC on 9 September 2017 and about 90 mm of rainfall between 01:30 UTC and 03:30 UTC on 10 September 2017), causing the overflow of Rio Maggiore and leading to tragic events and great damage. In a previous paper, we have analyzed the flash flood that occurred near Catania on 21 February 2013 [15], in which the extraordinarily intense rainfall (50 mm in 1 h) caused serious damages to the local infrastructures and economy. In the case of Catania, there was no severe weather warning issued by the Italian Department of Civil Protection (DPC), partially due to low quantitative precipitation forecast by NWP. Conversely, the case study discussed in this work was predicted by the most common NWP models, although the timing and location of the precipitation event occurred over Livorno were slightly mispredicted. This will be evident when comparing the rain rate derived by the Italian Weather Radar Network (WRN) with the maps of accumulated rainfall forecast by weather research forecasting (WRF) model. Thus, even if the DCP issued severe weather warnings, unfortunately, the tragedy was not avoided. In particular, the warning color code was high ("orange"), but not at its maximum ("red"), which is necessary to notify citizen evacuation alert. Moreover, the code was orange for many areas in central Italy, in particular, in Tuscany, but there was a lack of a correct prediction of the areas that would have been hit by heavy rainfall as well as a distinction between high level areas of hydraulic hazard ("red zones") and areas for which there was no high risk. The missing red alert caused several arguments between the local authorities and the DPC. The "alert system" developed by the local administration for highly dangerous situations had not been triggered. The study does not want to investigate further the abovementioned issues, which

have been briefly illustrated just to provide an overview of the Livorno tragic event. The precipitation of 9 and 10 September 2017 in central Italy was forecast by many NWP models and monitored by WRN, and, for this reason, it might seem superfluous to deepen the case study by using remote sensing techniques. As will be explained later, for this case, the NWP forecast missed important information and, in particular, the location and the time of the extreme event, which are crucial for the prevention and securing of the population. The availability of near real-time observations characterized by high spatial and temporal resolution is fundamental to provide continuous monitoring of events in emergency situations. To this aim, rain rate measurements derived from WRN provide support to NWP forecast. However, WRN-derived rain rates are not always available because radar networks do not supply global coverage in contrast to remote sensing techniques. In order to get around lacks and inadequacies in the radar and gauges coverage, a recent study [16] proposes an algorithm that blends the radar, gauges, and satellite quantitative precipitation estimation over areas immediately outside the effective radar coverage. Moreover, the importance of the fusion between WRN and satellite observations in supporting now-casting is investigated in [17], where measurements from the Chinese operational geostationary satellite Fengyun-2E, providing cloud top information at frequent intervals (<1 h), and ground-based weather radars, have been combined to conduct a study of the isolated convection initiation over central eastern China. A different approach has been followed in this study, the application of three satellite-based techniques to the Livorno extreme rainfall event will be shown with the purpose of investigating their usefulness in supplying information about cloud classification and in characterizing rainy areas when WRN information is missing. This information could be very helpful for real-time monitoring of extreme events and for supporting forecasters in warning decisions. In order to better understand the usefulness of satellite-based observations and to have a wide overview of the data contribution to the monitoring of extreme events, the case study is analyzed also by using WRF-accumulated rain maps and the WRN-derived rain rate maps. The applied satellite-based techniques were developed at the National Research Council (CNR) of Italy. They use both passive microwave (PMW) observations from Low-Earth Orbiting (LEO) satellite and infrared (IR)/visible (VIS) observations from Geosynchronous Earth Orbit (GEO) satellite to achieve information about the rain rate/class and cloud classification.

The PMW observations yield precipitation estimates more directly than IR/VIS, since microwave radiation is able to penetrate precipitating cloud [18,19], and it is less affected by cloud droplets than the IR/VIS radiation. However, the spatial resolution of the PMW observations from LEO satellite is relatively low: e.g., the dimensions of the instantaneous field of view (IFOV) of the Atmospheric Sounding Unit-A (AMSU-A) [20] is approximately (48 $\times$ 48) km^2 at the sub-satellite point (SSP), and (80 $\times$ 150) km^2 at the edge of scan line. Similarly, for AMSU-B and microwave humidity sounder (MHS) [21], the IFOVs are approximately (16 $\times$ 16) km^2 at the SSP and (27 $\times$ 50) km^2 at the edge of scan line. For this reason, and for the low LEO satellites' temporal resolution, PMW observations are not optimal for the continuous monitoring of both large-scale and small-scale precipitation events. Conversely, IR/VIS measurements do not provide direct measurements of precipitation. On the other hand, they are characterized by higher spatial and temporal resolution than PMW observations: e.g., the spatial resolution of the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) is about (5 $\times$ 7) km^2 at mid-latitudes with a temporal resolution of 15 min, which better suit the requirements for the continuous monitoring of extreme events. The studies [22–29] are among those that use PMW observations to monitor extreme precipitation events. IR/VIS observations have been used in many techniques proposed to determine rainfall rate; among the others, there are [30–32]. The first works combining IR and PMW observations to determine rainfall information are those proposed by [33,34]. Later, [35] proposed a mixed GEO-IR/LEO-PMW technique for the retrieval of rain rates (RR) measurements, which was taken as a role model by [36–38].

The precipitation evolving technique (PET) [39] and the rain class evaluation from infrared and visible observations (RainCEIV) [40] technique, exploited in this paper, combine the advantages of geosynchronous IR/VIS and polar PMW observations to get continuous monitoring of the

Livorno extreme event, and, to obtain more accurate RR estimates than derived by using IR/VIS observations only. In addition, the cloud classification mask coupling of statistical and physics methods (C-MACSP) [15,40,41] is also used to provide information about the cloud cover with particular regard to the evolution of convective clouds. The peculiarities of these techniques, as well as the WRF model, are summarized in Section 2, while the synoptic analysis of the event, the results obtained by applying C-MACSP, PET, and RainCEIV to the Livorno case study and the WRF forecast of the event, are shown in Sections 3.1–3.3, respectively. These results are discussed in Section 4. Finally, Section 5 draws the conclusions. Appendix A lists the acronyms used in this work.

2. Materials and Methods

This paragraph describes the remote sensing techniques and the WRF model used for the analysis of the Livorno case study. In addition, the National Weather Radar Network of Italy is also described, whose RR estimates are considered to evaluate the performance of the satellite techniques.

A detailed overview of C-MACSP [15,41,42], RainCEIV [40], and PET [15,39] is proposed in [43]. All of the three remote sensing techniques exploit the high temporal resolution (15 min) of the MSG-SEVIRI [44] to provide a continuous monitoring of the extreme events. Table 1 lists the data used in input by C-MACSP, RainCEIV, and PET, and the modules used for the implementation of each technique. In particular, C-MACSP uses MSG-SEVIRI IR/VIS observations for cloud classification, while RainCEIV and PET combine MSG-SEVIRI observations and PMW-RR maps, at their original spatial and temporal resolution, to give RR classes and RR maps at MSG-SEVIRI spatial and temporal resolution. The PMW RR maps, used for the training phase of RainCEIV and C-MACSP and to initialize PET, were obtained by the operational precipitation estimation at microwave frequencies (OPEMW) [45,46]. These PMW-based algorithms use AMSU-B or MHS observations [47,48]. The AMSU-B/MHS revisit time is once/twice a day at mid-latitudes, while the dimension of their IFOV is about (16×16) km^2 at nadir, and increase up about (27×50) km^2 at the scan-line edge. By contrast, the IR/VIS MSG-SEVIRI spatial sampling is 3 km at the sub-satellite point (along the equator) and increases up to ~11 km at high latitudes. The availability of RR maps at MSG-SEVIRI spatial and temporal resolution is fundamental in extreme situations when continuous monitoring is crucial for alert purposes. PET and RainCEIV handle PMW RR maps and IR/VIS observations differently to obtain information on RR. PET uses the MSG-SEVIRI IR brightness temperature differences between 6.2 μm water vapor channel and the 10.8 μm window channel to propagate forward in time and space the last available RR map. RainCEIV uses the RR maps re-gridded on the MSG-SEVIRI grid to build the training dataset for the K-nearest neighbor mean (K-NNM). K-NMM classifier works with spectral and textural features determined in MSG-SEVIRI IR/VIS images to associate each MSG-SEVIRI pixel to either non-rainy (RR < 0.5 mm/h) or rainy class (moderate-rain when $0.5 \leq$ RR ≤ 4 mm/h, heavy-to-very-heavy-rain when RR > 4 mm/h). As RainCEIV, C-MACSP use tonal and textural features properly chosen by the Fisher criterion [49,50] as input for a non-parametric supervised classifier that classifies each pixel as clear, low/middle cloud, high optically thin cloud, high optically thick cloud, and cloud having a high probability to be convective cloud. The case study is analyzed also considering the RR values derived from the Italian WRN which is coordinated by DPC [51,52] in collaboration with research centers, the air traffic control service (ENAV), the regional authorities, and the meteorological service of the Italian Air Force (CNMCA). The network counts 10 C-band weather radars managed by the regional authorities, 2 C-band weather radars managed by ENAV, 6 C-band weather radars, and 2 X-band polarimetric radars belonging to DPC. The surface rate intensity (SRI, corresponding to RR in mm/h) is retrieved applying a reflectivity-rainfall (Z-R) relationship to the reflectivity values at the lowest level of the corrected radar volume. SRI values are compared with PET-RR values and RainCEIV-RR classes. To get a fair comparison, the SRI are collocated into the MSG-SEVIRI grid. The original SRI product is allocated on a grid of 1400×1400 km^2 with a spatial and temporal resolution of about 1 km and 15 min, respectively. For this case study, SRI are available every 30 min. The collocation process of the

SRI values into the MSG-SEVIRI grid was implemented as described in [40]. It is important to keep in mind that, due to the MSG-SEVIRI scan principle, the size of the MSG-SEVIRI spatial resolution is not the same all over the disc [53]. In detail, in the center of Italy, the MSG-SEVIRI footprint size is about 4 km (east–west extent) by 5–6 km (north–south extent).

Table 1. Overview of the remote sensing techniques based on satellite observations considered in this study. (All the acronyms are listed in Appendix A.)

Technique	Data Used in Input	Technique Modules	Output
C-MACSP	IR/VIS observations from MSG-SEVIRI	K-NN classifier and physics thresholds tests	Cloud classification
PET	MSG-SEVIRI-IR/VIS observations and RR maps from AMSU-B/MHS PMW observations	Morphing and calibration modules	Rain rate
RainCEIV	C-MACSP maps, MSG-SEVIRI-IR/VIS observations and RR maps from AMSU-B/MHS PMW observations	K-NNM classifier and physics thresholds tests	Rain classes

SRI values are compared with PET-RR values and RainCEIV-RR classes. To get a fair comparison, the SRI are collocated into the MSG-SEVIRI grid. The original SRI product is allocated on a grid of 1400 $\times$ 1400 km^2, with a spatial and temporal resolution of about 1 km and 15 min, respectively. For this case study, SRI are available every 30 min. The collocation process of the SRI values into the MSG-SEVIRI grid was implemented as described in [40]. It is important to keep in mind that, due to the MSG-SEVIRI scan principle, the size of the MSG-SEVIRI spatial resolution is not the same all over the disc [53]. In detail, in the center of Italy, the MSG-SEVIRI footprint size is about 4 km (east–west extent) by 5–6 km (north–south extent). The size of each pixel has been determined individually. By looking in North–South direction (Figure 10 in [53]) and in East–West direction (Figure 11 in [53]), it is evident that in the area of interest the major variability of the pixel dimensions is along N–S direction. The size of each MSG-SEVIRI pixel has been determined by assuming that the geographical coordinates refer to the central point of the pixel and that the dimension of each MSG-SEVIRI pixel is the average distance between the MSG-SEVIRI pixel under consideration and the 4 closest MSG-SEVIRI pixels. To summarize the procedure adopted for the collocation of radar samples into the MSG-SEVIRI grid [40], the radar samples totally enclosed into the MSG-SEVIRI pixels will be denominated RS samples. When the percentage of the rainy or non-rainy RS samples is higher than 80%, the SEVIRI pixel is considered for the validation. In particular, when the comparison has been carried out with PET RR results, the corresponding radar RR is the average of the RS RR values. Moreover, the rain class corresponding to the radar RR is also considered for comparison with RainCEIV.

The precipitation maps produced by the WRF model forecast have also been used in this study. The WRF-ARW model [54] is a numerical weather prediction system resulting from a joint effort of different research institutes coordinated by the National Center for Atmospheric Research (NCAR, http://www.wrf-model.org). It has been developed both for research and operational purposes and it has been used over a wide range of scales, from climate studies to large eddy simulations. It solves the fully compressible, non-hydrostatic Euler equations, using the terrain-following, hydrostatic-pressure vertical coordinate with vertical grid stretching. It is based on time-split integration, using a 2nd- or 3rd-order Runge–Kutta scheme with smaller time step for acoustic and gravity-wave modes. The horizontal grid is staggered Arakawa-C. A variety of schemes are provided for the model physics: microphysics schemes ranging from simplified physics suitable for idealized studies to sophisticated mixed-phase physics suitable for process studies and NWP. Advanced Research WRF (ARW) version 3.9.1 has been used for the simulation of the event. The WRF has been configured on a domain consisting of 350 $\times$ 380 grid points, and covering Italy and part of the western Mediterranean Sea with a horizontal resolution of 3 km. The European Centre for Medium-Range Weather Forecasts (ECMWF) 40-vertical-level forecasts at 0.125 degrees, with a temporal pass of 6 h, were used to initialize the WRF simulation. The run started at 9 September 00:00 UTC and provided a 36 h forecast.

At this resolution, the precipitation is explicitly computed (no cumulus scheme), hydrometeors are calculated using the aerosol-aware Thompson Scheme [55], the Mellor–Yamada–Janjic one-dimensional prognostic turbulent kinetic energy scheme with local vertical mixing [56,57] for the planetary boundary layer (PBL) parameterization is used, and the RRTMG [58] for the longwave and the shortwave radiation scheme.

3. Results: Livorno Case Study (9 and 10 September 2017)

This section shows the mesoscale analysis (Section 3.1) and the results obtained by C-MACSP, PET, RainCEIV (Section 3.2), and WRF (Section 3.3) for the Livorno case under study.

3.1. Mesoscale Analysis

The extraordinary intense rainfall hitting central Italy and reaching its brunt over Livorno city on 9 and 10 September 2017, originated from a large area of low pressure extending from Scandinavia to the western Mediterranean basin. This is clearly visible from the mean sea level pressure and 500 hPa geopotential at 00:00 UTC 10 September 2017, reported in Figure 1a. The low pressure was associated to a cold front moving eastward, which advected low level moist air toward the northwest of Italy, as visible by the thermal, wind, and equivalent temperature (Θ_e) fields in Figure 1b,c. Figure 1a–c show the large cloud system expanding from the West Mediterranean Basin to central/northern Italy, and identify the presence of the warm conveyor belt (WCB), responsible of a structured flux of warm moist air (Figure 1b, values of $\Theta_e > 320$ K) running in front of the cold front (Figure 1b, 20 ms^{-1} at 700 hPa). The main convective activity was in the warm air ahead of the front and in the narrow cold frontal rain-band as a consequence of the WCB dynamics that is able to generate very intense prefrontal storms [59]. In detail, on 9 September at 00 UTC, a shallow cyclone of 1009 hPa developed over the Balearic Islands and moved eastward. Several mesoscale convective systems (MCS) developed in the north Tyrrhenian coasts, due to the strong potential vorticity (PV) anomaly and the warm temperature advection shown in Figure 1b,d. The movement of the shallow low-pressure area toward the Gulf of Genoa was enhanced by the PV anomaly (Figure 1d) and supported the strengthening of the trough, which reached a value of 999 hPa at 00:00 UTC of 10 September (Figure 1a). The deep trough caused the advection of moist air over the Tyrrhenian coasts and increased the potential instability in the lower atmosphere, allowing for the development of deep convection in the northwest of Italy. This severe event ended due to the weakening and to the rapid southward movement of the cold front.

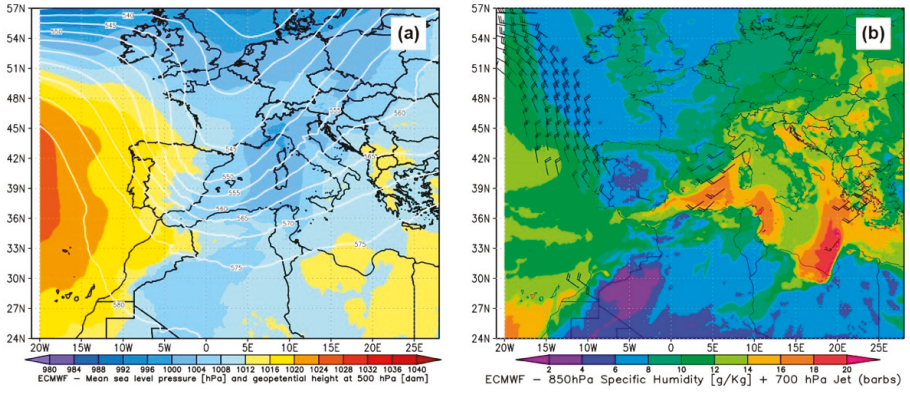


Figure 1. *Cont.*

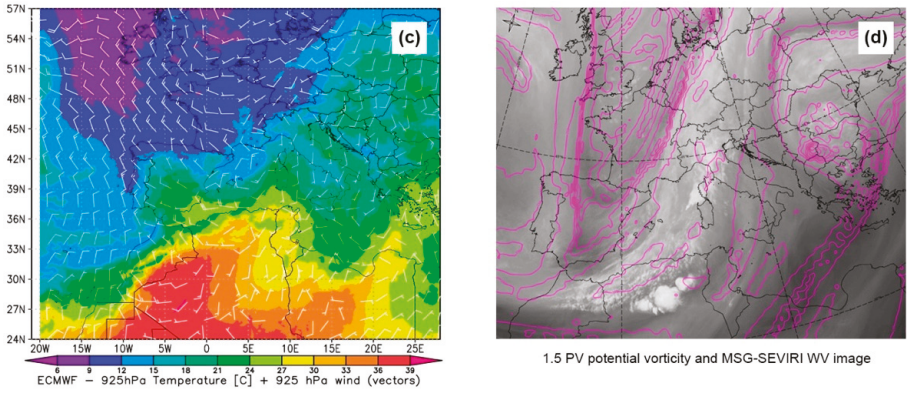


Figure 1. (**a**) Synoptic analysis at 00:00 UTC, 10 September 2017: European Centre for Medium-Range Weather Forecasts (ECMWF) 500 hPa geopotential height (white lines) and ECMWF mean sea level pressure; (**b**) Synoptic analysis at 18:00 UTC, 9 September 2017: ECMWF 850 hPa specific humidity (g/kg) and horizontal wind (speed greater than 20 m/s) at 700 hPa by ECMWF (barbs); (**c**) Synoptic analysis at 18:00 UTC, 9 September 2017: ECMWF 925 hPa temperature (°C) and horizontal wind at 925 hPa by ECMWF (barbs); (**d**) Synoptic analysis at 18:00 UTC, 9 September 2017: ECMWF height in hPa, where the potential vorticity (PV) equals 1.5 PV units and water vapor channel 6.2 μm of MSG—source: https://www.eumetrain.org.

3.2. Satellite Observations

The results obtained by C-MACSP, PET, and RainCEIV are presented in this subsection. A qualitative analysis, against RR values from the Weather Radar Network (WRN) and the rain gauges located in the area, is reported. The rain gauge RR measurements are provided by the Laboratory for the Environmental Monitoring and Modelling (LAMMA) of the Tuscany region (http://www.lamma.rete.toscana.it/meteo/osservazioni-e-dati/dati-stazioni, accessed on 17 July 2018), the Functional Center of the Tuscan region (http://www.cfr.toscana.it, accessed on 17 July 2018), and by the DEWETRA system [60]. The Livorno heavy rainfall event was characterized by two main peaks of precipitation. The first (Rainfall Peak A) occurred on 9 September 2017 between 20:45 UTC and 22:00 UTC—about 60 mm of accumulated precipitation were measured by the rain gauge located in Livorno (43.54 latitude north, 10.31 longitude east).

The second (Rainfall Peak B) occurred on 10 September 2017 between 01:30 UTC and 03:30 UTC, for which the same rain gauge measured about 90 mm of accumulated precipitation. Moreover, other rainfall peaks were recorded on 9 September 2017 in some stations near Livorno, such as Bocca d'Arno (about 25 km from Livorno) and Stagno (about 10 km from Livorno), where 64 mm between 22:00 UTC and 23:00 UTC, and 52 mm between 19:00 UTC and 20:00 UTC, respectively, were revealed. The rain gauge in Valle Benedetta (about 11 km from Livorno) detected an hourly maximum greater than 90 mm on 10 September 2017 between 01:00 UTC and 03:00 UTC. The rainfall areal distribution corresponding to Rainfall Peak A is shown in Figure 2 from the perspective of the RR maps measured by WRN. The 20:30 UTC RR map shows a heavy precipitation cell that is approaching Livorno from the northwest, and precisely from Bocca d'Arno, where about 30 mm of rainfall were recorded between 20:30 UTC and 21:30 UTC. The 21:00 UTC and 21:30 UTC WRN RR maps show intense precipitation on Livorno, which dissipated around 22:00 UTC while moving north of Livorno. The sequences of time co-located MACSP, RainCEIV, and PET maps are shown in Figures 3–5, respectively. The first PET/MACSP/RainCEIV map is related to the temporally closest AMSU-B/MHS overpass in the area of interest (approximately 20:41 UTC). The PMW-RR field, used to initialize PET, was obtained by applying the OPEMW algorithm [33] to the AMSU-B/MHS observations acquired at 20:41 UTC.

The PET and RainCEIV detect heavy rainfall (RR > 4 mm/h) between 20:41 UTC and 21:26 UTC in agreement with the WRN RR maps, but it is evident that RainCEIV overestimates rainy areas over the Tyrrhenian sea and northwestern Italy. Specifically, C-MACSP (Figure 3) correctly detects optically thick high cloud located over and near Livorno and convective clouds North of Livorno from 20:41 UTC to 21:56 UTC. In correspondence with optically thick high clouds and convective clouds detected by C-MACSP, RainCEIV (Figure 4) associates heavy and moderate RR class to the pixels over and near Livorno from 20:42 UTC to 21:56 UTC. The rainy area identified by PET is less extended than that detected by RainCEIV, more similarly to the WRN RR maps. All the three satellite-based techniques agree with WRN in tracking the storm evolution. In particular, they correctly individuate the mature stage (from 20:41 UTC to 21:26 UTC) and the dissipating stage (from 21:56 UTC to 22:56 UTC) over Livorno. Moreover, according WRN RR maps, they correctly detect the temporal evolution of convective clouds and heavy precipitation at northwest Livorno, from approximately 22:00 UTC to 03:30 UTC. From 21:56 UTC to 22:56 UTC, C-MACSP does not detect convective activity over and near Livorno, and PET and RainCEIV detect no rainy and moderate rainy in the same area, according to C-MACSP and WRN RR maps. Heavy precipitation begins to approach Livorno from 23:30 UTC, and it reaches its maximum during Rainfall Peak B, as testified also by C-MACSP, RainCEIV, and PET, showing persistent convective clouds and heavy rainfall from 01:30 UTC to 03:00 UTC.

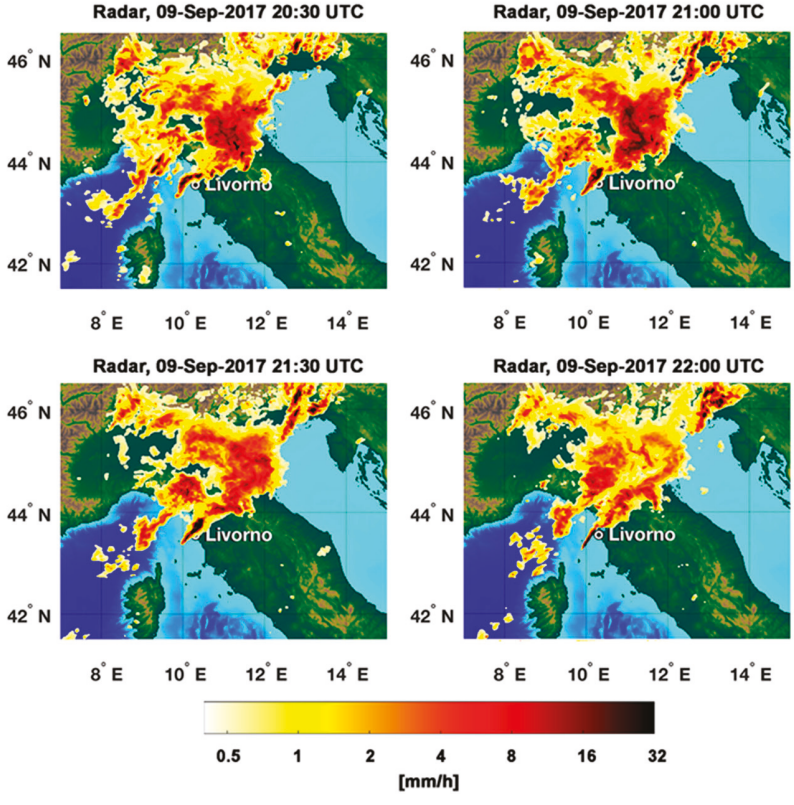


Figure 2. WRN (Weather Radar Network) rain rate maps for 9 September 2017 from 20:30 (**left**) to 22:00 (**right**).

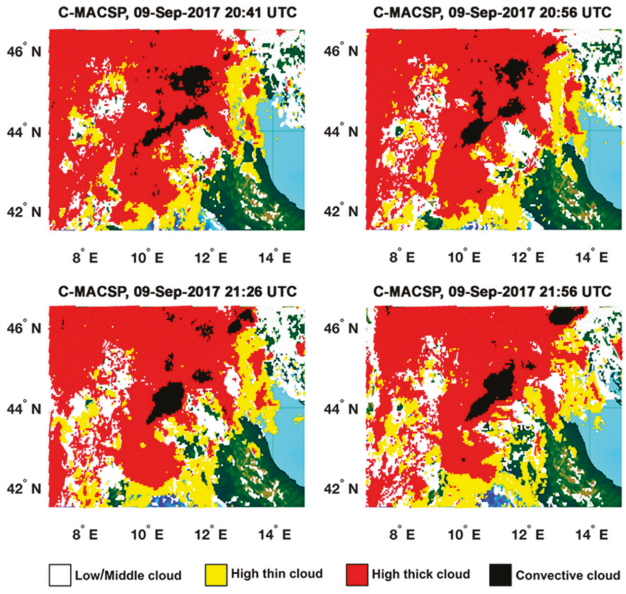


Figure 3. C-MACSP (cloud Classification–MAsk Coupling of Statistical and Physical methods) maps for 9 September 2017 from 20:26 UTC to 21:56 UTC (Rainfall Peak A).

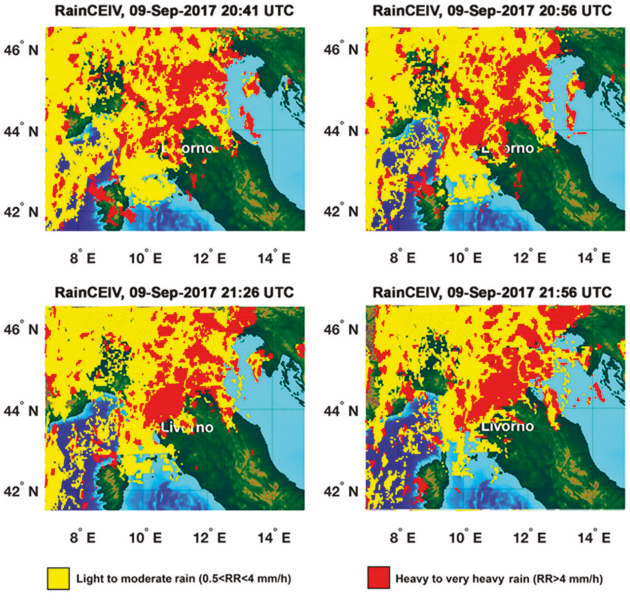


Figure 4. RainCEIV (Rain Class Evaluation from Infrared and Visible observations) maps for 9 September 2017 from 20:41 UTC to 21:56 UTC (Rainfall Peak A).

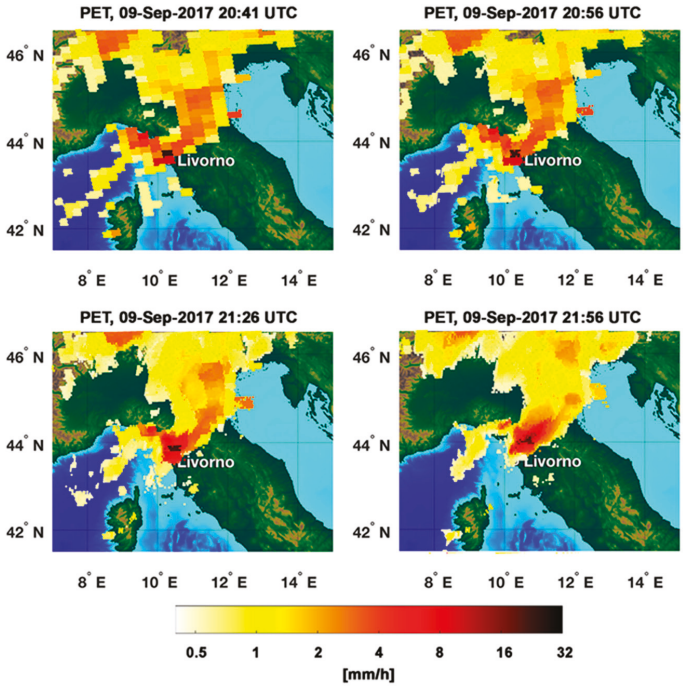


Figure 5. PET (Precipitation Evolving Technique) maps for 9 September 2017 from 20:41 UTC to 21:56 UTC (Rainfall Peak A). The spatial resolution appears coarser in the first PET RR map (20:41 UTC); this is due to the direct resampling of the 20:41 UTC PMW RR map into the MSG-SEVIRI grid without applying the PET modules. PET modules have been applied subsequently, giving better spatial resolution to the following images.

For Rainfall Peak B (from 01:30 UTC to 03:00 UTC of 10 September), the WNR RR, C_MACSP, RainCEIV, and PET RR maps are shown in Figures 6–9, respectively. In detail, WRN RR maps show heavy precipitation over and near Livorno from 01:30 UTC to 03:00 UTC of 10 September, and, in agreement with them, the C-MACSP detects convective clouds and high optically thick clouds over the same area, RainCEIV associates to the cloudy pixels detected by C-MACSP to the C_1 and, prevalently, to the C_2 class. Also, in the case of B Rainfall Peak, by a visual comparison with WRN RR maps, it is evident that RainCEIV overestimates the extent of rainy areas, while PET underestimates the rainy areas over the Tyrrhenian sea near Livorno in the 01:26 UTC and 01:56 UTC maps. In the following two maps (02:26 UTC and 02:56 UTC), the rainy areas are well detected by PET. The 02:26 UTC map for PET corresponds to the AMSU overpass, which has been used to initialize the PET; because of this, the PET performs better from 02:26 UTC to 02:56 UTC than in the previous maps. Figure 10 shows the temporal distribution of C_0, C_1, and C_2 RR classes (0, 1, and 2, respectively, in the left panel) and of WRN RR (right panel) in about (20 × 20) km^2 area surrounding Livorno. The WRN RR at 22:00 UTC is missing near Livorno; in addition, WRN RR maps are not available at 23:00 UTC of 9 September 2017, and at 01:00 UTC of 10 September 2017. In Figure 10, it is evident that WRN detects moderate rain near Livorno at 20:30 UTC of 9 September 2017 in agreement with RainCEIV, while PET detects heavy rain. WRN, PET, and RainCEIV detect heavy rain from 21:00 UTC to 21:30 UTC of 9 September. Moreover, WRN RR is higher than 4 mm h^{-1} from 23:00 UTC of 9 September 2017 to 03:30 UTC of 10 September 2017, in agreement with both PET and RainCEIV at 00:30 UTC, and from 02:30 UTC to

02:30 UTC of 10 September 2017. PET and RainCEIV agree in detecting moderate rain at 22:30 UTC and from 23:30 UTC of 9 September to 00:00 UTC of 10 September 2017.

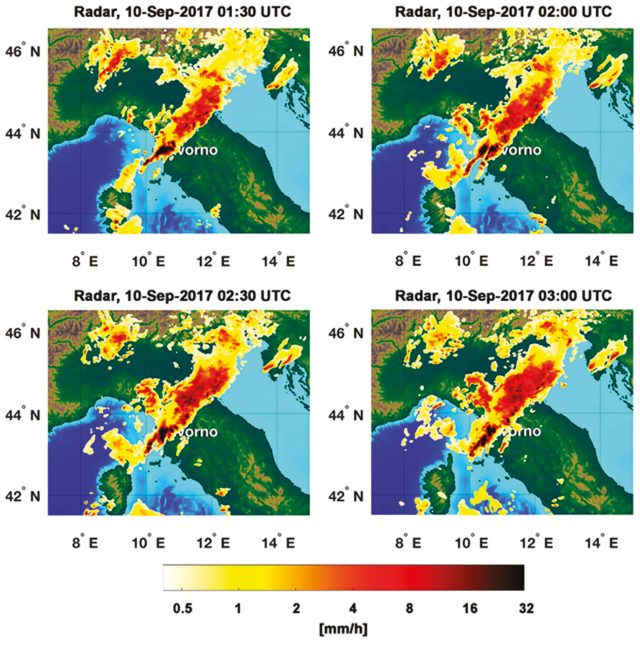


Figure 6. WRN RR maps for 10 September 2017 from 01:30 UTC to 03:00 UTC: Rainfall Peak B over Livorno.

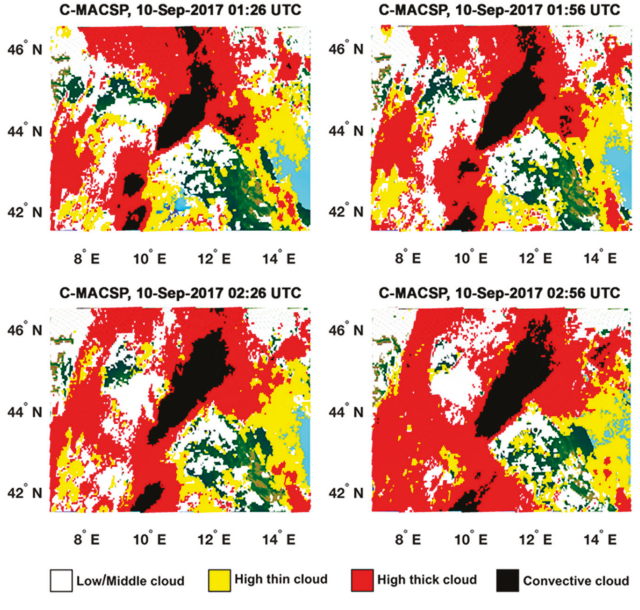


Figure 7. C-MACSP maps for 10 September 2017 from 01:26 UTC to 02:56 UTC (Rainfall Peak B).

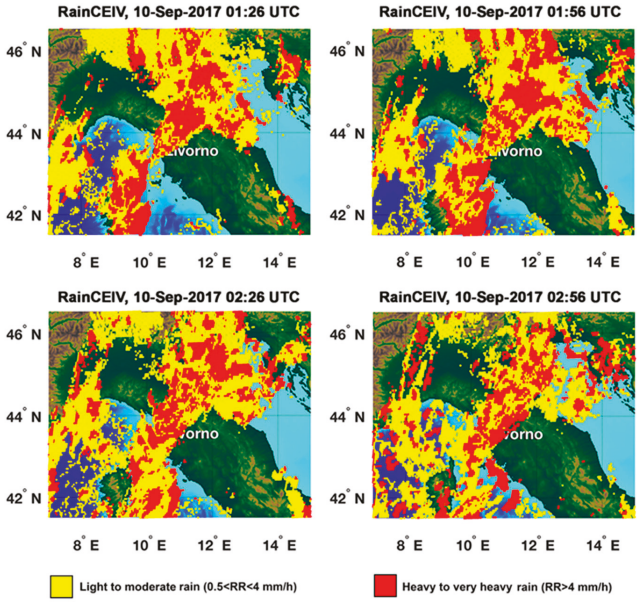


Figure 8. RainCEIV maps for 10 September 2017 from 01:26 UTC to 02:56 UTC (Rainfall Peak B).

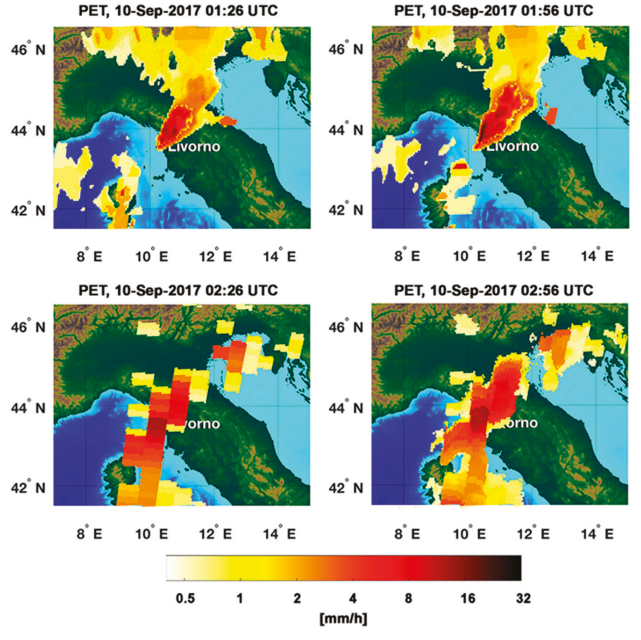


Figure 9. PET maps for 10 September 2017 from 01:26 UTC to 02:56 UTC (Rainfall Peak B). The coarser resolution of the 02:26 UTC is due to the fact that PET has been re-initialized by using the PMW RR information available at 02:26 UTC.

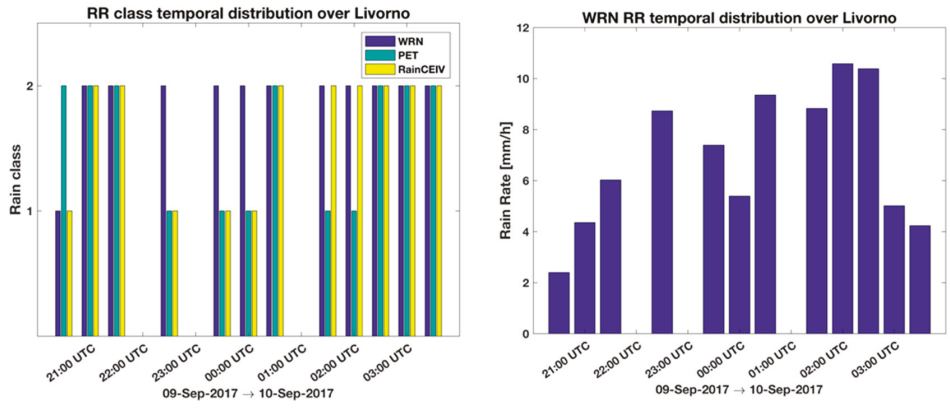


Figure 10. Temporal intervals distributions of WRN, PET, and RainCEIV RR classes (**left panel**) and WRN RR temporal distribution (**right panel**) over Livorno from 20:30 UTC of 9 September 2017 to 03:30 UTC of 10 September 2017.

3.3. WRF Analysis

The Rainfall Peak A has not been forecast by WRF, as it is evident comparing the 21:00 UTC and 22:00 UTC WRF rainfall accumulated maps (Figure 11) with the corresponding WRN maps (Figure 2). WRF forecasts heavy precipitation in the Ligurian Gulf near Genoa and La Spezia, but no precipitation has been forecast in proximity of Livorno. Figure 12 shows the WRF rainfall cumulated maps of 02:00 UTC and 03:00 UTC. The heavy rainfall that hit Livorno has not been forecast by WRF, which predicts heavy rain only near La Spezia (northwest of Livorno) while moderate rain is predicted on the Tyrrhenian sea southwest of Livorno at 02:00 UTC. Thus, near Livorno, WRF forecasts an hourly accumulated rainfall lower than 4 mm, and no precipitation in the subsequent hour. It is evident that the real-time availability of PET, RainCEIV, and C-MACSP information would have been very useful to observe and provide the continuous monitoring of the event not forecast by the WRF model.

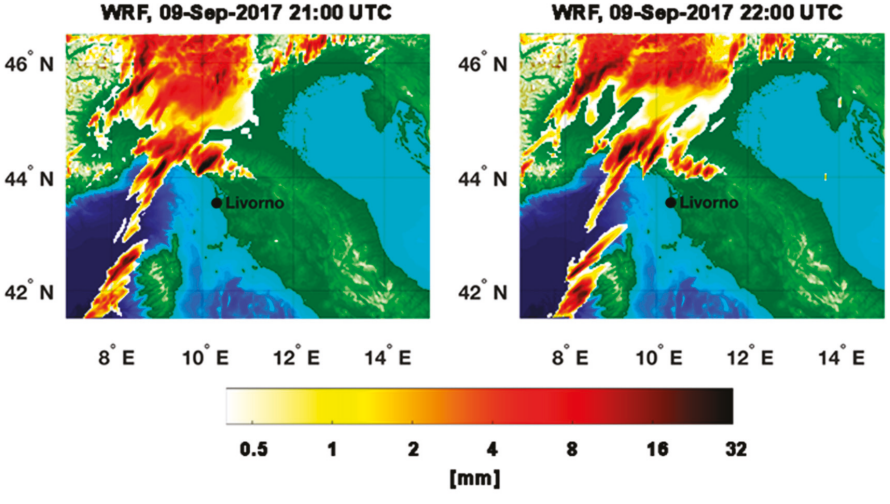


Figure 11. WRF (Weather Research and Forecasting) rainfall accumulated maps for 9 September 2017 at 21:00 UTC and 22:00 UTC.

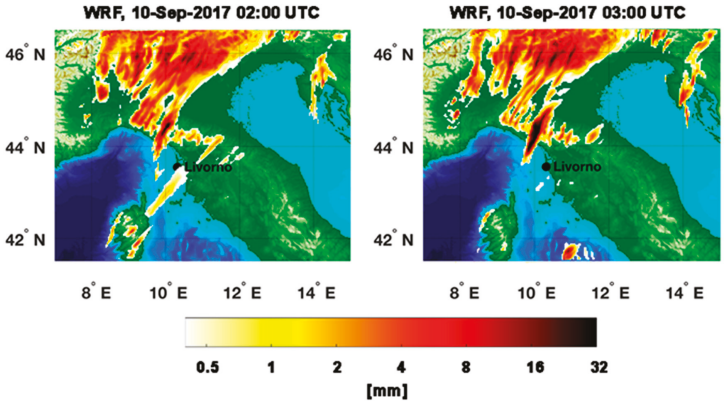


Figure 12. WRF rainfall accumulated maps for 10 September 2017 at 02:00 UTC and 03:00 UTC.

4. Discussion

A quantitative evaluation of PET and RainCEIV estimates is performed through a dichotomous statistical assessment. It is important to bear in mind that the PET and Radar RR maps, as well as the RainCEIV RR class map, refers to instantaneous RR values while the WRF RR maps refers to hourly accumulated rainfall. Due to this, we preferred to carry out the quantitative analysis only between instantaneous RR values derived from WRN and from PET and RainCEIV. The dichotomous assessment of PET and RainCEIV is proposed in order to highlight the different performances of the two techniques, and to individuate weakness and strengths and, thus, understand their effectiveness in supporting NWP models. This support could be very useful especially when radar information is not available due to technical problems or partial coverage, e.g., https://www.rainviewer.com/coverage.html (accessed on 2 August 2018). In such cases, the support of satellite-based techniques is vital to forecasters for providing support in managing local alerts. The PET and WRN RR values have been associated to the same RainCEIV RR classes so to have comparable statistics for RainCEIV and PET. Moreover, the radar RR values have been co-located with MSG-SEVIRI, as described in Section 2. In detail, pixels are associated to the class C_0, C_1, or C_2, as defined for RainCEIV [40], when the corresponding RR value satisfies the relation $0 \leq RR < 0.5$ mm h^{-1} (C_0, non-rainy class), $0.5 \leq RR < 4$ mm h^{-1} (C_1, light to moderate rain class) or $RR \geq 4$ mm h^{-1} (C_2, heavy rain class), respectively. This rain classification is arbitrary, and it is based on the observation of rain rate associated to stratiform rain and to convective rain at mid latitudes. In particular, the non-rainy class ranges from 0 mm h^{-1} to 0.1 or 0.5 mm h^{-1}, because these small values are within the instrument sensitivity and, thus, are included in the non-rainy class. The skill scores used for dichotomous assessment are accuracy, bias score, probability of detection (POD), false alarm ratio (FAR), and Heidke skill score (HSS), described in [61]. Tables 2 and 3 summarize the contingency values related to Rainfall Peak A, while the skill scores, calculated for all the classes considered together, as well as for C_1 and C_2 classes separately, are shown in Table 4 for both RainCEIV and PET statistical assessment. For this case study, it is evident that PET performs better than RainCEIV; in fact when classes C_1 and C_2 are considered together, the PET accuracy score (0.87) is higher than that related to RainCEIV (0.78) for all the rainy classes. The bias is slightly higher than 1 for PET (1.06), also showing high POD (0.85) and low FAR (0.20). These scores indicate that PET detects rainy areas with better performances and reduce the overestimation of rainy areas with respect to RainCEIV (bias = 1.28, FAR = 0.35). Although bias is higher for RainCEIV than for PET, due to higher false alarms number for RainCEIV (8391) than PET (3936), *POD* is high for both the algorithms (0.83 for RainCEIV, 0.85 for PET). *HSS* is relative to random forecast, and it is based on accuracy corrected by the number of hits that would be expected by chance. HSS can assume values between -1 and 1, and its perfect score is 1, while 0 indicates no skill. HSS values

for PET and RainCEIV are 0.75 and 0.55, respectively, demonstrating good performance of the two algorithms in detecting rainy areas and confirming the best performance of PET. The statistical scores related to the C_1 and C_2 classes have been also considered to further analyze the behavior of the two algorithms. In order to better explain the C_1 and C_2 statistical assessment, it is helpful to define the C_1 and C_2 contingency table elements. Specifically, for the dichotomous statistical assessment of C_1/C_2 class, the hits or correct negatives are the pixels for which both the PET/RainCEIV RR class, and the corresponding WRN co-located RR class are C_1/C_2 or C_0 classes, respectively. For the dichotomous statistical assessment of C_1 class, the misses are the pixels for which the WRN co-located RR class is C_1, and the corresponding PET/RainCEIV RR class is C_0, while the false alarms are the pixels for which the PET/RainCEIV RR class is C_1, while the corresponding class is C_0. Finally, for dichotomous statistical assessment of C_2 class, the misses are the pixels for which WRN co-located RR class is C_2, and the corresponding PET/RainCEIV RR class is C_0 or C_1, and the false alarms are the pixels for which PET/RainCEIV RR class is C_2, and the corresponding WRN co-located RR class is C_0 or C_1. The C_1/C_2 accuracy for RainCEIV is 0.71/0.87, which is lower than that of PET (0.85/0.91), consistently with the fact that both hits and correct negatives are higher for PET than for RainCEIV for both the classes. The RainCEIV tendency to overestimate rainy areas is confirmed also for the two classes separately; in fact, the RainCEIV FAR is 0.60 for C_1 and 0.51 for C_2 while the PET FAR is 0.30 and 0.05 for C_1 and C_2 class, respectively. Along with the higher FAR values, RainCEIV shows better performances in detecting rainy pixels belonging both to C_1 and C_2 class. RainCEIV skill in detecting C_2 rainy areas seems better (for Rainfall Peak A) than that of PET: C_2 bias and *POD* are 1.00 and 0.49 for RainCEIV while 0.30 and 0.28 for PET. This latter result is strongly influenced by the fact that RainCEIV C_2 hits are higher than the PET C_2 hits; on the contrary, C_2 correct negatives are smaller for RainCEIV than for PET, resulting in higher C_2 accuracy for PET (0.91) than for RainCEIV (0.87).

Table 2. Contingency table for the dichotomous statistical assessment of RainCEIV for Rainfall Peak A (from 20:30 UTC to 21:30 UTC of 9 September 2017).

		Radar-Derived Rain Rate Results		
		Rain	No rain	Marginal total
RainCEIV results	Rain	15,309	8391	23,700
	No rain	3150	25,710	28,860
	Marginal total	18,459	34,101	52,560

Table 3. Contingency table for the dichotomous statistical assessment of PET for Rainfall Peak A (from 20:30 UTC to 21:30 UTC of 9 September 2017).

		Radar-Derived Rain Rate Results		
		Rain	No rain	Marginal total
PET results	Rain	15,723	3936	19,659
	No rain	2736	30,165	32,901
	Marginal total	18,459	34,101	52,560

Table 4. Dichotomous statistical scores (RainCEIV/PET versus radar-derived rain rate measurements) for Rainfall Peak A (from 20:30 UTC to 21:30 UTC of 9 September 2017). The statistical scores are shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy rain (C_2).

Statistical Scores	**C_1 (RainCEIV)**	**C_1 (PET)**	**C_2 (RainCEIV)**	**C_2 (PET)**	**C_1, C_2 (RainCEIV)**	**C_1, C_2 (PET)**
Accuracy	0.71	0.85	0.87	0.91	0.78	0.87
Bias	1.79	1.17	1.00	0.30	1.28	1.06
POD	0.71	0.81	0.49	0.28	0.83	0.85
FAR	0.60	0.30	0.51	0.05	0.35	0.20
HSS	0.31	0.65	0.34	0.34	0.55	0.75

For Rainfall Peak B, Tables 5 and 6 summarize the contingency values and Table 7 shows the statistical assessment for all the classes, confirming the good performance of PET (accuracy = 0.78, bias = 0.94, POD = 0.66, FAR = 0.30, HSS = 0.51) for the case study analyzed. The RainCEIV statistical assessment for all the classes reveals accuracy (0.69) and HSS (0.37) skill to be smaller than the corresponding PET statistical scores, the bias (1.28) indicates a tendency in RainCEIV overestimating rainy areas, while POD (0.70) is slightly higher than the PET POD. As evident also by visual comparison of RainCEIV with Radar RR (Section 3.2), it is clear that RainCEIV overestimates rainy areas and, in particular, the C_1 rainy areas (bias = 1.56, FAR = 0.67). For RainCEIV C_2 class, bias (1.38) and FAR (0.64) are slightly better than the C_1 class corresponding values, and when compared with PET C_2 statistical scores, it is possible to note a much smaller FAR (0.17), but also a bias (0.83) that reveals the tendency of PET in underestimating C_2 rainy areas, while this is not the case for C_1 class (bias = 1.09, POD = 0.51). It is evident that, while RainCEIV tendency (overestimation of rainy areas) is the same for the two rainfall peaks, the PET statistical behavior is strongly affected by the fact that at 01:26 UTC and 01:56 UTC PET has been still moving forward the information that corresponds to the 20:42 UTC AMSU initial map. This means that the temporal distance from the PMW initialization data is higher than 4 h, while the PET statistical assessment in [39] shows the PET overall ability to propagate rain field at least for 2–3 h. The statistical assessment of RainCEIV and PET is slightly better for Rainfall Peak A than for Rainfall Peak B. This depends on the different developing stages of the storm related to the two peaks. In detail, the storm associated to Rainfall Peak A is composed by convective cells in initial and mature stage whose cloud-tops and the related sub-cloud rain estimations are determined with smaller overestimations than those characterizing the detection of convective cloud top and rain estimations associated with mature and dissipating stages. In fact, the statistical assessment of Rainfall Peak B is strongly influenced by the presence of a large convective activity starting at circa 22:00 UTC northwest Livorno; this is a mix of mature and dissipating stages because of its long duration. The presence of dissipating stage causes the convective cloud top detected by IR to appear larger and, consequently, the rain estimation in sub-cloud area is overestimated. This effect is more evident in RainCEIV than in PET because, in PET, it is mitigated by the direct use of PMW RR.

Table 5. Contingency table for the dichotomous statistical assessment of RainCEIV for Rainfall Peak B (from 00:30 UTC to 03:30 UTC of 10 September 2017).

		Radar-Derived Rain Rate Results		
		Rain	No rain	Marginal total
RainCEIV results	Rain	16,306	13,412	29,718
	No rain	6950	29,032	35,982
	Marginal total	23,256	42,444	65,700

Table 6. Contingency table for the dichotomous statistical assessment of PET for the Rainfall Peak B (from 00:30 UTC to 03:30 UTC of 10 September 2017).

		Radar-Derived Rain Rate Results		
		Rain	No rain	Marginal total
PET results	Rain	15,242	6574	21,816
	No rain	8014	35,870	43,884
	Marginal total	23,256	42,444	65,700

Table 7. Dichotomous statistical scores (RainCEIV/PET versus radar-derived RR measurements) for Rainfall Peak B (from 00:30 UTC to 03:30 UTC of 10 September 2017). The statistical scores are shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy rain (C_2).

Statistical Scores	$C_{1\ (RainCEIV)}$	$C_{1\ (PET)}$	$C_{2\ (RainCEIV)}$	$C_{2\ (PET)}$	C_1, $C_{2\ (RainCEIV)}$	C_1, $C_{2\ (PET)}$
Accuracy	0.64	0.73	0.82	0.95	0.69	0.78
Bias	1.56	1.09	1.38	0.83	1.28	0.94
POD	0.52	0.51	0.49	0.69	0.70	0.66
FAR	0.67	0.53	0.64	0.17	0.45	0.30
HSS	0.16	0.30	0.22	0.69	0.37	0.51

Finally, Tables 8 and 9 show the contingency values related to the whole storm duration (from 21:30 UTC of 10 September 2017 to 03:30 UTC of 10 September 2017) while Table 10 shows the skill scores, calculated for all the classes considered together, as well as for C_1 and C_2 classes separately, for both RainCEIV and PET statistical assessment. The PET statistical results for all the classes (accuracy = 0.83, POD = 0.76, FAR = 0.26, HSS = 0.62, bias = 1.03) are slightly better than the RainCEIV ones (accuracy = 0.74, POD = 0.77, FAR = 0.42, HSS = 0.46, bias = 1.33), and confirm the RainCEIV tendency in overestimating rainy areas. The statistical scores related to both the algorithms are comparable with the ones obtained by [14], who investigated the reliability of several PMW-based precipitation algorithms [62–65], obtaining POD values between 0.60–0.76 and FAR between 0.28–0.45, and with the statistical assessment of CMORPH [12], which shows $0.20 < POD < 0.80$, $0.15 < FAR < 0.45$, and $0.30 < HSS < 0.62$.

Table 8. Contingency table for the dichotomous statistical assessment of PET for the whole storm duration (from 20:30 UTC of 9 September 2017 to 03:30 UTC of 10 September 2017).

		Radar-Derived Rain Rate Results		
		Rain	No rain	Marginal total
PET results	Rain	40,142	14,219	54,361
	No rain	12,611	90,708	103,319
	Marginal total	52,753	104,927	157,680

Table 9. Contingency table for the dichotomous statistical assessment of RainCEIV for the whole storm duration (from 20:30 UTC of 10 September 2017 to 03:30 UTC of 10 September 2017).

		Radar-Derived Rain Rate Results		
		Rain	No rain	Marginal total
RainCEIV results	Rain	40,866	29,448	70,314
	No rain	11,887	75,479	87,366
	Marginal total	52,753	104,927	157,680

Table 10. Dichotomous statistical scores (RainCEIV/PET versus radar -derived RR measurements) for the whole storm duration (from 20:30 UTC of 9 September to 03:30 UTC of 10 September 2017). The statistical scores are shown for both rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy rain (C_2).

Statistical Scores	$C_{1\ (RainCEIV)}$	$C_{1\ (PET)}$	$C_{2\ (RainCEIV)}$	$C_{2\ (PET)}$	C_1, $C_{2\ (RainCEIV)}$	C_1, $C_{2\ (PET)}$
Accuracy	0.69	0.80	0.85	0.93	0.74	0.83
Bias	1.69	1.16	1.33	0.57	1.33	1.03
POD	0.63	0.69	0.49	0.49	0.77	0.76
FAR	0.62	0.41	0.63	0.13	0.42	0.26
HSS	0.27	0.49	0.25	0.56	0.46	0.62

The proposed satellite-based techniques show their ability in detecting rainy areas and in characterizing them by using cloud type information from C_MACSP. RainCEIV and PET, providing near-real-time estimations of RR and RR classes at the MSG-SEVIRI temporal and spatial resolution (i.e., about 15 min and (5×7) km^2 at mid-latitude), show different performance for both rainfall peaks. Compared to WRN RR, we found better agreement with PET than with RainCEIV, which is attributable to the different way the two techniques deal with the information from PMW RR. In fact, PET propagates forward, in space and in time, the last available PMW RR map that gives direct information on the event to analyze. This influences the slightly worse PET performance in Rainfall Peak B. In fact, it is a consequence of the different delay from the last available PMW observation (PMW delay) used to initialize PET in the two Peaks. In detail, PMW delay is about 1 h for Rainfall Peak A and about 4 h for the 01:26 UTC and 01:56 UTC PET maps of Rainfall Peak B, producing a higher RR overestimation in Peak B than in Peak A. On the contrary, no direct information on the actual rainfall distribution is provided to RainCEIV. The only PMW information that RainCEIV uses is within the training dataset built on a series of PMW RR maps from previous rainfall events. However, in spite of its worse statistical scores, RainCEIV is able to give information on rain class without temporal and spatial limitations (within the observed Earth's disk), while PET produces RR maps for only a few hours after the last available PMW observations, and only on the region around the PMW swath observations. From the qualitative analysis and statistical assessment, it is evident that the three techniques are complementary in supporting the WRF model forecasting, especially for real-time continuous monitoring of convective events. In particular, PET can provide information about RR for about three hours from the last PMW RR map available on the area of interest, while RainCEIV and C_MACSP can give information about RR class and cloud type with no limitation of space and time within the observed Earth's disk. When available, PET RR information is preferable to the RainCEIV RR class information. On the other hand, RainCEIV RR class information is useful when PET is not available. For the case study examined in this paper, the information from the three satellite-based techniques could appear redundant as a support for the NWP forecast, because of the availability of WRN-derived RR. In any case, the aim of the analysis presented is to highlight the effectiveness of the satellite techniques as a valid support for NWP forecast, especially when there is no WRN information.

5. Conclusions

The prediction of the intense precipitation that hit central Italy on 9 and 10 September 2017 did not help to avoid the tragic events and serious damage to the local infrastructure in the city of Livorno, although it was forecast by several NWP models. Precipitation was intense everywhere in Tuscany; DPC alerted the whole region at the same level, though Livorno was the only city heavily affected, prevalently for reasons related to the urban development near Rio Maggiore. In this paper, the event has been analyzed by means of remote sensing techniques, to investigate their effectiveness in continuous monitoring of extreme events and their utility in supporting NWP forecast. To this end, precipitation maps from the WRF model forecast have also been analyzed. All the three remote sensing techniques used in this study exploit the high spatial and temporal resolution of MSG-SEVIRI, in order to give information about the cloud classification (C-MACSP), the RR class (RainCEIV), and the RR values (PET). These techniques use the RR information retrieved by the PMW observations differently, and this affects their performance. In detail, PET exploits the latest PMW observations directly, and it provides RR map values every 15 min on the area limited to the PMW swath and for a few hours. On the contrary, RainCEIV and C-MACSP indirectly use the PMW RR maps to build the training dataset, which is not updated with the most recent information on the raining area distribution. RainCEIV RR class maps and C-MACSP cloud classification maps are available every 15 min without temporal and spatial limitation within the Earth's disk, observed by GEO orbit. It is evident that C-MACSP, PET, and RainCEIV are complementary for the real-time monitoring of extreme events. In fact, C-MACSP gives information about cloud cover and cloud type, while PET and RainCEIV give RR quantitative and qualitative information, respectively. Due to the abovementioned PET temporal

and spatial limitations, PET RR values are not always available and, in this case, RainCEIV RR class information, although less efficient than PET quantitative information, is very useful in order to detect and characterize rainy areas. The techniques agree with WRN-RR maps in detecting rainy areas, as confirmed by the qualitative analysis for both A and B rainfall peak. Moreover, the statistical scores determined for the Rainfall Peaks A/B and for the entire of the event's duration (from 20:30 UTC of 9 September to 03:30 UTC of 10 September 2017) indicate good performances of PET and RainCEIV (for PET: accuracy = 0.82, bias = 1.03, POD = 0.76, FAR = 0.26, HSS = 0.62; for RainCEIV: accuracy = 0.74, bias = 1.33, POD = 0.77, FAR = 0.41, HSS = 0.46). In fact, the bias and FAR are higher for RainCEIV (1.33 and 0.41, respectively) than for PET (bias = 1.03, FAR = 0.26), denoting a tendency to overestimate the extent of rainy areas for RainCEIV. The WRF hourly accumulated precipitation maps do not detect heavy precipitation in Livorno and its surroundings in any of the rainfall peaks analyzed. The three satellite techniques could be very useful for the real-time monitoring purposes, especially when the forecast is not correct and the WRN data are not available as support for forecasters. Clearly, MACSP, PET, and RainCEIV cannot replace the NWP maps in foreseeing the precipitation area, but they can be very useful to characterize the area and observe the severe weather in real-time, especially for the cases in which WRN RR information is missing. In particular, RainCEIV RR class maps are to be used when the PET RR maps are not available, whereas C-MACSP gives information about cloud cover and can be useful to detect areas with high probability for deep convection.

Author Contributions: Conceptualization, E.R., F.D.P., S.G., D.C. and F.R.; Data curation, E.R., F.D.P., S.G., A.C., D.C., D.G., E.G., S.L., E.R., F.R. and M.V.; Methodology, E.R., F.D.P., and S.G.; Validation, E.R., F.D.P., S.G. and D.C.; Visualization, D.C. and S.T.N.; Writing—original draft, E.R., and D.C.

Funding: This work has been financed by the Italian Ministry of Economic Development (MISE) in the framework of the SolarCloud project, contract No. B01/0771/04/X24.

Acknowledgments: The authors acknowledge the Italian Department of Civil Protection for providing data from the national radar network and the rain gauge network.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A. Acronyms

Table A1. Acronyms list.

Acronyms	Meaning
AMSU	Advanced Technology Microwave Sounder
C-MACSP	cloud Classification-MAsk Coupling
DCP	Department of Civil Protection
FAR	False Alarm Ratio
HSS	Heidke Skill Score
IR/VIS	Infrared/Visible
K-NNM	K-Nearest Neighbor Mean classifier
MSC	Mesoscale Convective System
MHS	Microwave Humidity Sounder
MSG/SEVIRI	Meteosat Second Generation/Spinning Enhanced Visible and Infrared Imager
NWP	Numerical Weather Prediction
OPEMW	Operational Precipitation Estimation at Microwave frequencies
PET	Precipitation Evolving Technique
PMW	Passive Microwave Measurements
POD	Probability Of Detection
PV	Potential Vorticity
RainCEIV	Rain Class Evaluation from Infrared and Visible observations
RR	Rain Rate
WCB	Warm Conveyor Belt
WRF	Weather Research Forecasting
WRN	Weather Radar Network
WV	Water Vapor

References

1. Pregnolato, M.; Sean, A.F.; Wilkinson, M.; Dawson, R.J. The impact of flooding on road transport: A depth-disruption function. *Transp. Res.* **2017**, *55*, 67–81. [CrossRef]
2. Mecikalski, J.R.; Berendes, T.A.; Feltz, W.F.; Bedka, K.M.; Bedka, S.T.; Murray, J.J.; Wimmers, A.J.; Minnis, P.; Johnson, D.B.; Haggerty, J.; et al. Aviation Applications for Satellite-Based Observations of Cloud Properties, Convection Initiation, In-Flight Icing, Turbulence, and Volcanic Ash. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 1589–1607. [CrossRef]
3. Murray, J.J. Aviation weather applications of Earth Science Enter- prise data. *Earth Obs. Mag.* **2002**, *11*, 27–30.
4. Cuo, L.; Pagano, T.C.; Wang, Q.J. A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting. *J. Hydrometeorol.* **2011**, *12*, 713–728. [CrossRef]
5. Merk, D.; Zinner, T. Detection of convective initiation using Meteosat SEVIRI: Implementation in and verification with the tracking and nowcasting algorithm Cb-TRAM. *Atmos. Meas. Tech.* **2013**, *6*, 1903–1918. [CrossRef]
6. Zinner, T.; Mannstein, H.; Tafferner, A. Cb-TRAM: Tracking and monitoring severe convection from onset over rapid devel- opment to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorol. Atmos. Phys.* **2008**, *101*, 191–210. [CrossRef]
7. Zinner, T.; Forster, C.; de Coning, E.; Betz, H.-D. Validation of the Meteosat storm detection and nowcasting system Cb-TRAM with lightning network data—Europe and South Africa. *Atmos. Meas. Tech.* **2013**, *6*, 1567–1583. [CrossRef]
8. Mecikalski, J.R.; Bedka, K.M. Forecasting Convective Initiation by Monitoring the Evolution of Moving Cumulus in Day- time GOES Imagery. *Mon. Weather Rev.* **2006**, *134*, 49–78. [CrossRef]
9. Li, Z.; Li, J.; Wang, P.; Lim, A.; Li, J.; Schmit, T.J.; Atlas, R.; Boukabara, S.A.; Hoffman, R.F. Value-added Impact of Geostationary Hyperspectral Infrared Sounders on Local Severe Storm Forecasts via a Quick Regional OSSE. *Adv. Atmos. Sci.* **2018**, *35*, 1217–1230. [CrossRef]
10. Ai, Y.; Li, J.; Shi, W.; Schmit, T.J.; Cao, C.; Li, W. Deep convective cloud characterizations from both broadband imager and hyperspectral infrared sounder measurements. *J. Geophys. Res. Atmos.* **2017**, *122*, 1700–1712. [CrossRef]
11. Neto, C.P.S.; Barbosa, H.A.; Beneti, C.A.A. A method for convective storm detection using satellite data. *Atmósfera* **2016**, *29*, 343–358. [CrossRef]
12. 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]
13. Joyce, R.J.; Xie, P. Kalman filter–based cmorph. *J. Hydrometeorol.* **2011**, *12*, 1547–1563. [CrossRef]
14. Panegrossi, G.; Casella, D.; Dietrich, S.; Marra, A.C.; Sanò, 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. Select. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2733–2753. [CrossRef]
15. Di Paola, F.; Ricciardelli, E.; Cimini, D.; Romano, F.; Viggiano, M.; Cuomo, V. Analysis of Catania Flash Flood Case Study by Using Combined Microwave and Infrared Technique. *J. Hydrometeorol.* **2014**, *15*, 1989–1998. [CrossRef]
16. He, Y.; Zhang, Y.; Kuligowski, R.; Cifelli, R.; Kitzmiller, D. Incorporating Satellite Precipitation Estimates into a Radar-Gauge Multi-Sensor Precipitation Estimation Algorithm. *Remote Sens.* **2018**, *10*, 106. [CrossRef]
17. Huang, Y.; Meng, Z.; Li, J.; Li, W.; Bai, L.; Zhang, M.; Wang, X. Distribution and Variability of Satellite-Derived Signals of Isolated Convection Initiation Events over Central Eastern China. *J. Geophys. Res. Atmos.* **2017**, *122*, 21. [CrossRef]
18. Weng, F.; Grody, N.C. Retrieval of ice cloud parameters using a microwave imaging radiometer. *J. Atmos. Sci.* **2000**, *57*, 1069–1081. [CrossRef]
19. 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]
20. Grody, N.; Zhao, J.; Ferraro, R.; Weng, F.; Boers, R. Determination of precipitable water and cloud liquid water over oceans from the NOAA 15 Advanced Microwave Sounding Unit. *J. Geophys. Res.* **2001**, *106*. [CrossRef]

21. Bonsignori, R. The Microwave Humidity Sounder (MHS): In-orbit performance assessment. *Proc. SPIE* **2007**, *6744*, 6744. [CrossRef]
22. Ferraro, R.R.; Weng, F.; Grody, N.C.; Zhao, L.; Meng, H.; Kongoli, C.; Pellegrino, P.; Qiu, S.; Dean, C. NOAA operational hydrological products derived from the Advanced Microwave Sounding Unit. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1036–1049. [CrossRef]
23. Wilcox, E.M.; Donner, L.J. The frequency of extreme rain events in satellite rain-rate estimates and an atmospheric general circulation model. *J. Clim.* **2007**, *20*, 53–69. [CrossRef]
24. Wardah, T.; Abu Bakar, S.H.; Bardossy, A.; Maznorizan, M. Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting. *J. Hydrol.* **2008**, *356*, 283–298. [CrossRef]
25. Kummerow, C.D.; Ringerud, S.; Crook, J.; Randel, D.; Berg, W. An observationally generated a priori database for microwave rainfall retrievals. *J. Atmos. Ocean. Technol.* **2010**, *28*, 113–130. [CrossRef]
26. Bennartz, R.; Schroeder, M. Convective activity over Africa and the tropical Atlantic inferred from 20 years of geostationary Meteosat infrared observations. *J. Clim.* **2012**, *25*, 156–169. [CrossRef]
27. Wu, H.; Adler, R.F.; Hong, Y.; Tian, Y.; Policelli, F. Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. *J. Hydrometeorol.* **2012**, *13*, 1268–1284. [CrossRef]
28. Casella, D.; Dietrich, S.; di Paola, F.; Formenton, M.; Mugnai, A.; Porcù, F.; Sanò, P. PM-GCD—A combined IR–MW satellite technique for frequent retrieval of heavy precipitation. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 231–240. [CrossRef]
29. Sanò, P.; Panegrossi, G.; Casella, D.; Di Paola, F.; Milani, L.; Mugnai, A.; Petracca, M.; Dietrich, S. The passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for AMSU/MHS observations: Description and application to European case studies. *Atmos. Meas. Tech.* **2015**, *8*, 837–857. [CrossRef]
30. Adler, R.F.; Negri, A.J. A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteorol.* **1988**, *27*, 30–51. [CrossRef]
31. Adler, R.F.; Markus, M.J.; Fenn, D.D. Detection of severe Midwest thunderstorms using geosynchronous satellite data. *Mon. Weather Rev.* **1985**, *113*, 769–781. [CrossRef]
32. Wu, R.; Weinman, J.A.; Chin, R.T. Determination of rainfall rates from GOES satellite images by a pattern recognition technique. *J. Atmos. Ocean. Technol.* **1985**, *2*, 314–330. [CrossRef]
33. 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]
34. Jobard, I.; Desbois, M. Satellite estimation of the tropical precipitation using the Meteosat and SSM/I data. *Atmos. Res.* **1994**, *34*, 285–298. [CrossRef]
35. Turk, J.F.; 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 Environment*; Pampaloni, P., Ed.; VSP Intern. Sci. Publisher: Utrecht, The Netherlands, 2000; pp. 353–363.
36. Kidd, C.; Kniveton, D.R.; Todd, M.C.; Bellerby, T.J. Satellite rainfall estimation using a combined passive microwave and infrared algorithm. *J. Hydrometeorol.* **2003**, *4*, 1088–1104. [CrossRef]
37. Marzano, F.S.; Palmacci, M.; Cimini, D.; Giuliani, G.; Turk, J.F. 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]
38. Heinemann, T.; Lattanzio, A.; Roveda, F. The Eumetsat multi- sensor precipitation estimate (MPE). In Proceedings of the Second International Precipitation Working Group (IPWG) Meeting, Madrid, Spain, 23–27 September 2002.
39. Di Paola, F.; Casella, D.; Dietrich, S.; Mugnai, A.; Ricciardelli, E.; Romano, F.; Sanoò, P. Combined MW-IR Precipita- tion Evolving Technique (PET) of convective rain fields. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 3557–3570. [CrossRef]
40. Ricciardelli, E.; Cimini, D.; Di Paola, F.; Romano, F.; Viggiano, M.A. Statistical approach for rain intensity differentiation using Meteosat second generation-spinning enhanced visible and infrared imager observations. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2559–2576. [CrossRef]
41. Ricciardelli, E.; Romano, F.; Cuomo, V. Physical and statistical approaches for cloud identification using Meteosat Second Generation-Spinning Enhanced Visible and Infrared Imager Data. *Remote Sens. Environ.* **2008**, *112*, 2741–2760. [CrossRef]

42. Nilo, S.T.; Romano, F.; Cermak, J.; Cimini, D.; Ricciardelli, E.; Cersosimo, A.; Di Paola, F.; Gallucci, D.; Gentile, S.; Geraldi, E.; et al. Fog Detection Based on Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager High Resolution Visible Channel. *Remote Sens.* **2018**, *10*, 541. [CrossRef]
43. Ricciardelli, E.; Di Paola, F.; Cimini, D.; Romano, F.; Viggiano, M. Satellite Remote Sensing for clouds and precipitation. In *Advances in Watershed Hydrology*; Water Resource Publications, LLC: Littleton, CO, USA, 2015; pp. 69–82, ISBN 154-196, 978-1-887201-85-8.
44. Schmetz, J.; Pili, P.; Tjemkes, S.; Just, D.; Kerkmann, J.; Rota, S.; Ratier, A. An Introduction to Meteosat Second Generation (MSG). *Bull. Am. Meteorol. Soc.* **2002**, *83*, 977–992. [CrossRef]
45. Di Tomaso, E.; Romano, F.; Cuomo, V. Rainfall estimation from satellite passive microwave observations in the range 89 GHz to 190 GHz. *J. Geophys. Res.* **2009**, *114*, D18203. [CrossRef]
46. Cimini, D.; Romano, F.; Ricciardelli, E.; Di Paola, F.; Viggiano, M.; Marzano, F.S.; Colaiuda, V.; Picciotti, E.; Vulpiani, G.; Cuomo, V. Validation of satellite OPEMW precipitation product with ground-based weather radar and rain gauge networks. *Atmos. Meas. Tech.* **2013**, *6*, 3181–3196. [CrossRef]
47. Kramer, H.J. *Observation of the Earth and Its Environment*; Springer: Berlin, Germany, 2002; p. 1510.
48. Kleespies, T.J.; Watts, P. Comparison of simulated radiances, jacobians and linear error analysis for the Microwave Humidity Sounder and the Advanced Microwave Sounding Unit-B. *Q. J. R. Meteorol. Soc.* **2007**, *132*, 3001–3010. [CrossRef]
49. Ebert, E. A pattern recognition technique for distinguishing surface and cloud types in the polar regions. *J. Clim. Appl. Meteorol.* **1987**, *26*, 1412–1427. [CrossRef]
50. Parikh, J. A comparative Study of Cloud Classification Techniques. *Remote Sens. Environ.* **1977**, *6*, 67–81. [CrossRef]
51. Vulpiani, G.; Pagliara, P.; Negri, M.; Rossi, L.; Gioia, A.; Giordano, P.; Alberoni, P.P.; Cremonini, R.; Ferraris, L.; Marzano, F.S. The Italian radar network within the national early-warning system for multi-risks management. In *Fifth European Conference on Radar in Meteorology and Hydrology (ERAD 2008)*; Finnish Meteorological Institute: Helsinki, Finland, 2008.
52. Vulpiani, G.; Montopoli, M.; Delli Passeri, L.; 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]
53. MSG Level 1.5 Image Data Format Description. 2017. Available online: https://www.eumetsat.int/website/home/Data/TechnicalDocuments/index.html (accessed on 12 July 2018).
54. Skamarock, W.C.; Joseph, B.K.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Duda, M.G.; Huang, X.Y.; Wang, W.; Powers, J.G. *A Description of the Advanced Research WRF Version 3*; NCAR: Boulder, CO, USA, 2008.
55. Thompson, G.; Eidhammer, T. A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.* **2014**, *71*, 3636–3658. [CrossRef]
56. Janjic, Z.I.; Gerrity, J.P., Jr.; Nickovic, S. An alternative approach to nonhydrostatic modeling. *Mon. Weather Rev.* **2001**, *129*, 1164–1178. [CrossRef]
57. Mellor, G.L.; Yamada, T. Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.* **1982**, *20*, 851–875. [CrossRef]
58. 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.* **2008**, *113*, D13103. [CrossRef]
59. Catto, J.L.; Jakob, C.; Nicholls, N. Can the CMIP5 models represent winter frontal precipitation? *Geophys. Res. Lett.* **2015**. [CrossRef]
60. Comellas, A.; Llasat, M.C.; Molini, L.; Parodi, A.; Siccardi, F. Severe rainfall prediction systems for civil protection purposes. In Proceedings of the 12th EGU Plinius Conference on Mediterranean Storms Plinius Conference Abstracts, Corfu, Greece, 1–4 September 2010.
61. Ebert, E. Forecast Verification Issues, Methods and FAQ. Available online: http://www.cawcr.gov.au/projects/verification/ (accessed on 12 July 2018).
62. Casella, D.; Panegrossi, G.; Sano, P.; Dietrich, S.; Mugnai, A.; Smith, E.A.; Tripoli, G.J.; Formenton, M.; Di Paola, F.; Leung, W.-Y.H.; et al. Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite overcoming database profile selection ambiguity by consideration of meteorological control on microphysics. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4650–4671. [CrossRef]

63. Sanò, P.; Casella, D.; Mugnai, A.; Schiavon, G.; Smith, E.A.; Tripoli, G.J. Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive microwave measurements: Part 1. Algorithm description and testing. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4119–4143. [CrossRef]
64. Mugnai, A.; Smith, E.A.; Tripoli, G.J.; Bizzarri, B.; Casella, D.; Dietrich, S.; Di Paola, F.; Panegrossi, G.; Sanò, P. CDRD and PNPR satellite passive microwave precipitation retrieval algorithms: EuroTRMM/EURAINSAT origins and H-SAF operations. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 887–912. [CrossRef]
65. Global Precipitation Measurement (GPM) Mission Algorithm Theoretical Basis Document Version 1.4 (GPROF2014 Conical Version Running at the PPS). Available online: http://rain.atmos.colostate.edu/ATBD/ATBD_GPM_Aug1_2014.pdf (accessed on 3 September 2018).

remote sensing

MDPI

Article

Evaluation of TRMM/GPM Blended Daily Products over Brazil

José Roberto Rozante [1,*], Daniel A. Vila [1], Júlio Barboza Chiquetto [2], Alex de A. Fernandes [1] and Débora Souza Alvim [3]

1 Center for Weather Forecast and Climate Studies, National Institute for Space Research, Rodovia Presidente Dutra, Km 40, SP-RJ-CEP, Cachoeira Paulista, SP 12630-000, Brazil; daniel.vila@inpe.br (D.A.V.); alex.fernandes@inpe.br (A.d.A.F.)
2 Department of Geography, University of Sao Paulo (USP), São Paulo, SP 05508-080, Brazil; julio22@usp.br
3 Center of Natural and Human Sciences (CCNH), Federal University of ABC (UFABC), Santo André, SP 09210-580, Brazil; debora.alvim@ufabc.edu.br
* Correspondence: roberto.rozante@inpe.br; Tel.: +55-12-3208-8400

Received: 1 March 2018; Accepted: 11 April 2018; Published: 6 June 2018

Abstract: The precipitation estimates from the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (named TMPA and TMPA-RT for the near real-time version) are widely used both in research and in operational forecasting. However, they will be discontinued soon. The products from the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (IMERG) and The Global Satellite Mapping of Precipitation (GSMaP) are analyzed as potential replacements for TMPA products. The objective of this study is to assess whether the IMERG and/or GSMaP products can properly replace TMPA in several regions with different precipitation regimes within Brazil. The validation study was conducted during the period from 1st of April, 2014 to the 28th of February, 2017 (1065 days), using daily accumulated rain gauge stations over Brazil. Six regions were considered for this study: five according to the precipitation regime, plus another one for the entire Brazilian territory. IMERG-Final, TMPA-V7 and GSMaP-Gauges were the selected versions of those algorithms for this validation study, which include a bias adjustment with monthly (IMERG and TMPA) and daily (GSMaP) gauge accumulations, because they are widely used in the user's community. Results indicate similar behavior for IMERG and TMPA products, showing that they overestimate precipitation, while GSMaP tend to slightly underestimate the amount of rainfall in most of the analyzed regions. The exception is the northeastern coast of Brazil, where all products underestimate the daily rainfall accumulations. For all analyzed regions, GSMaP and IMERG products present a better performance compared to TMPA products; therefore, they could be suitable replacements for the TMPA. This is particularly important for hydrological forecasting in small river basins, since those products present a finer spatial and temporal resolution compared with TMPA.

Keywords: precipitation; satellite; GPM

1. Introduction

Knowledge of the spatial and temporal distribution of precipitation is of key importance for planning a wide range of socio-economic activities such as agriculture, livestock grazing, energy generation, etc. The availability of accurate and consistent precipitation data is then paramount for a proper assessment of such activities. However, traditional rain gauge measurements are relatively scarce and poorly distributed over the surface of the globe, particularly over remote areas or in developing countries. In the last three decades, satellite-derived precipitation estimate products have been developed using multi-satellites and multi-sensors. Such products provide an effective way of

estimating precipitation data in areas where measurements are scarce, such as deserts [1], forests [2] and oceans [3]. Accordingly, they have been widely used in research and applications worldwide [4–7].

The first approaches to employ remote sensing techniques for estimating precipitation were performed during the 70s. During that time, satellite images were not digitized. Barret [8] was one of the pioneers in developing a method for estimating monthly precipitation, using the visible channel. This method was called "cloud index", and was later improved by Follansbee [9], which included estimations of daily precipitation rates. Almost two decades later, at the end of 1997, the Tropical Rainfall Measuring Mission (TRMM) [10] was launched jointly by NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency), aiming to improve precipitation estimates in tropical and subtropical regions. One of the most successful products generated from this mission is the Multi-Satellite Precipitation Analysis research version (TMPA) and the real-time version of the same product (TMPA-RT). Besides combining precipitation estimates derived from several satellites, the TMPA algorithm [4,11,12] is also able to incorporate observed precipitation data [13]. Over the last decade, those retrievals have been greatly improved with different versions of those products. For this study, the version of TMPA launched in 2012 (version 7 or TMPA-V7 hereafter) will be used.

The TMPA product has been used in scientific research and operational activities which lead to outstanding socio-economic gains, such as studies of extreme precipitation events [14–16], forecasting of natural disasters [17,18], water resources management planning [19], performance of numerical models [20,21], among others. Due to this, TMPA validation studies have been performed in several regions around the globe, and those results show great agreement between TMPA products and surface data [17,22–26]. Nevertheless, particularly over Brazilian territory, systematic bias are still observed for some precipitation regimes associated with shallow convection systems near the coast of northeastern Brazil (underestimation of precipitation) [27,28]. Conversely, in Southern Brazil (close to the border with Argentina and Paraguay), this algorithm overestimates the observed rainfall [20,26]. According to Laing and Fritsch [29], one of the largest and most active mesoscale convective complexes (MCCs) in the world is observed in this region.

In the beginning of 2014, the Global Precipitation Measurement (GPM) mission was launched to improve global estimates of precipitation and snow in low and mid latitudes. Moreover, GPM is a natural replacement of the successful TRMM mission [30]. The precipitation estimate algorithm created to replace TMPA is the IMERG (Integrated Multi-satellite Retrievals for GPM), made publicly available in the beginning of 2015 in the NASA portal [12]. This suite of products, namely IMERG-Early, IMERG-Late and IMERG-Final, is considered the next-generation of satellite-derived precipitation products, since it brings together resources from the existing: (1) TMPA [4], (2) CMORPH (Climate Prediction Center Morphing) [5], and (3) Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks—Cloud Classification System (PERSIANN-CCS) [6,31]. Because they are a recent effort, the products generated from GPM require urgent evaluation in order to be available for use, compared to other precipitation estimate products.

On the other hand, the Global Satellite Mapping of Precipitation (GSMaP) product, developed by a consortium of Japanese institutions and operated by JAXA [7,32], also offers a global coverage of rainfall with several versions: the real time version (GSMaP_NRT), the standard version (GSMaP_MVK) and the standard version with gauge correction (GSMaP-Gauge) [33]. The basic idea of the GSMaP algorithm is to find the optimal precipitation for which the brightness temperatures (TBs) calculated by the radiative-transfer model (RTM) fit best with observed TBs [7].

In view of the need to evaluate the accuracy of the new suite of GPM products in different regions of the globe, this study aims to assess the IMERG-Final and GSMaP-Gauge retrievals, in order to replace the TMPA research version in several regions with different precipitation regimes in Brazil. This paper is organized as follows: Section 2 provides details of the study area, the criteria used for the division of sub-areas according to precipitation regimes, and the datasets and statistics used; Sections 3

and 4 present the main results of this research and a discussion. The main conclusions are provided in Section 5.

2. Materials and Methods

2.1. Area of Interest and Characterization of the Precipitation Regimes

Brazil, due to its continental dimensions (8,515,759 km^2 of territorial area), features a great diversity of landscapes, topography, biodiversity and climates, as well as of precipitation regimes. In order to evaluate the effectiveness of satellite-derived precipitation products in the country, it is necessary to contextualize the main precipitation regimes in Brazil. In order to do that, 18 years of precipitation data (1998–2016) from the MERGE product [34] were used. This product combines daily precipitation from rain gauge stations (see observed data section) with the TMPA_RT product. According to this study, the MERGE product has proven to be a valuable analysis tool for model evaluation, outperforming gauge analysis for those regions with low rain gauges density.

Figure 1 shows the spatial distribution of the precipitation climatology over Brazil based on MERGE data. This climatology was calculated for the whole country and divided into grid cells with 2×2 degrees. This spatial distribution fits very well with the study of Reboita et al. [35] based on gauge analysis. Considering these results and the Figure 3 of Reboita et al., five different precipitation regimes were identified within Brazilian territory. Region R1, located in southern Brazil, presents year-round well distributed precipitation, and high total precipitation: 1050–1750 mm/year. The main systems influencing this region are cold fronts, mesoscale convective complexes [36], the South Atlantic Convergence Zone (ZCAS) [37] and the low-level jet stream [38]. Region R2, with a clear monsoon regime [39], covers most of the Brazilian territory, where the accumulated precipitation is higher during summer (DJF) and lower during winter time (JJA). Region R3, the driest region of the Brazilian territory, is located inland over northeastern Brazil and also presents a maximum of precipitation during summer and a minimum during winter, with totals between 200 and 500 mm/year. The main systems influencing this region are the Intertropical Convergence zone (ITCZ) and upper level cyclonic vortexes [40]. Region R4 (located in the northeastern coast of Brazil) presents maximum precipitation in winter and minimum in summer. The main systems are the ITCZ, tropical mesoscale convective systems, the Trade Winds, upper level cyclonic vortexes, easterly waves and sea breeze circulation [41]. Region R5 comprises the northern Amazonian region; the main influencing systems are the ITCZ, the tropical squall lines [42] and the trade winds.

2.2. Observed Data

The 24-h accumulation of rain gauge data used in this study are received daily, in near real time, by the Centre for Weather Forecasting and Climatic Studies of Brazil (CPTEC/INPE), and cover the period 12 UTC to 12 UTC. Main sources of precipitation data are composed of the global telecommunication system (GTS), the automatic platform for collecting data (PCDs), and the regional meteorological centers in Brazil. Most surface stations are located in eastern Brazil, near the coast. Towards the center of the continent, the network density decreases sharply. A quality control of gauge data is performed in two stages: the first is objective and the second is subjective. The data is verified in real time, that is, at the time of storage in database systems. At this stage, validity, consistency (internal and spatial) and control (temporal and climatological) checks are performed. In the validity check, acceptable values must belong to a predefined interval (or be within a tolerance limit). Internal consistency performs validations of variables focused on internal relationships, considering a single weather station. The limits of the variables (precipitation, in this case) are reevaluated in a spatial consistency process, considering different climatic regions. This process also compares variables from different meteorological stations (for example, the precipitation of a station is compared with the precipitation of other stations within a distance of up to 25 km). The temporal control verifies the differences of the variables over a given period (for example, precipitation is acceptable if the

difference of its current value in relation to the previous value does not exceed 175 mm). At the end of the first stage, the data are classified according to quality, by means of a descriptor, whose values are: suspect or correct. In the subjective stage, verification is performed by a meteorologist, who evaluates variables which have been identified as suspicious in the first stage. After the subjective analysis, suspect variables can be classified as valid or invalid. This process was carried out over 35 months, from the 1st of April, 2014, to the 28th of February, 2017; on average, 3400 daily observations were used in this study.

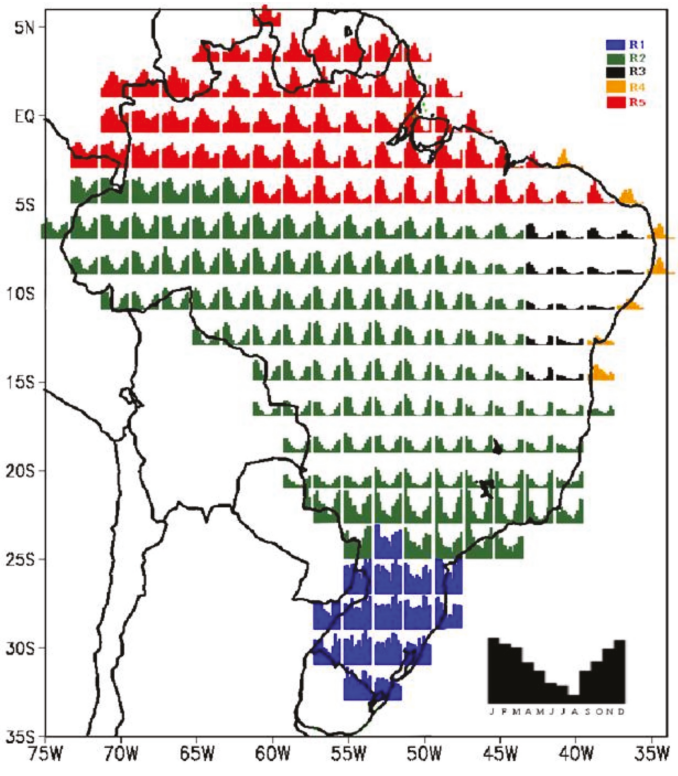


Figure 1. Spatial distribution of precipitation climatology (1998–2016) based on MERGE data [34] for the five identified regions, for each grid box of approximately 2 degrees.

2.3. TMPA Products

TMPA is responsible for two products: the research V7 version (TMPA), and the real-time version (TMPA-RT). Both are used extensively by a large community of users [11]. The products combine the estimates of precipitation from several satellite sensors. The biggest difference between them is that TMPA (TMPA-V7 hereafter) incorporates rainfall monthly data from the global rain gauge network (GPCC, Global Precipitation Climate Center) to scale the final product. These products are developed with a temporal resolution of 3 h, and a horizontal resolution of 0.25°, covering most of the globe (between latitudes 50°N–S). TMPA products were accumulated over 24 h, according to guidelines from the World Meteorological Organization. This is done in order to standardize the time of synoptic observations around the world, according to universal time: the total precipitation from 12 UTC of a given day up to 12 UTC of the following day is used. TMPA data were obtained via ftp from the US

National Aeronautics and Space Administration (NASA) (https://disc2.gesdisc.eosdis.nasa.gov/data). In order to use the final version of each product, TMPA-V7 will be used in this comparison.

2.4. IMERG Product

The Global Precipitation Measurement Core Observatory Satellite (GPM), launched on the 27th of February, 2014, aims to provide the next generation of precipitation products, continuing with the first-rate products provided by TMPA. Its algorithm, The Integrated Multisatellite Retrievals for GPM (IMERG), is similar to the algorithm of the TMPA products. It was built to calibrate, combine and interpolate satellite-derived precipitations (microwave, infrared) and worldwide observed data. IMERG is executed in near real-time for operational purposes and with two months' delay for the research version. It provides two near real-time precipitation estimates data options: Early and Late. Early provides a quick estimate, with a 4 h lag, taking into account only data which has been collected at that moment. Late has a 12-h lag (after more data has arrived), and is therefore obviously more precise. For the IMERG research product, estimates are combined with monthly observed data (similar procedure applied for TMPA-V7), and made available two months later (called IMERG-Final, hereafter IMERG-F).

IMERG-F product (version V05) used in this study were obtained from ftp://arthurhou.pps.eosdis.nasa.gov/gpmallversions/V05/with a temporal resolution of 30 min and $0.1^\circ \times 0.1^\circ$ of horizontal resolution. IMERG-F covers most of the globe: all surface area between latitudes 60°N and 60°S, corresponding to 87% of the Earth's surface. According to WMO guidelines, IMERG-F data have also been accumulated over 24 h, as well as TMPA-V7 data. The high spatial and temporal resolution, together with the expressive area of operation, make IMERG-F a potentially valuable tool for the scientific community. In this study, IMERG-F data (the research version) have been used for comparisons with other algorithms. More details on the GPM products and the IMERG algorithm can be found in [12].

2.5. GSMaP Product

Development of the precipitation map algorithm, including microwave radiometer/sounder algorithms, has been continued in cooperation with the members of Global Satellite Mapping of Precipitation (GSMaP) project [32] in Japan. Since the GSMaP project targeted the production of the "best" precipitation estimates, and they did not consider real-time operation and/or data availability, JAXA has developed and has operated a global rainfall map production system in near-real-time since October 2008; hourly and 0.1-degree resolution binary data and images are available via the internet (http://sharaku.eorc.jaxa.jp/GSMaP/) four hours after observation. Core algorithms of the system are based on those provided by the GSMaP project; microwave radiometer rainfall retrieval algorithm [7], microwave sounder rainfall retrieval algorithm [43], microwave imager/sounder rainfall retrieval algorithm [44], microwave-infrared (IR) merged algorithm [45] and Gauge calibrated rainfall algorithm [33].

GSMaP-Gauge (hereafter GSMaP-G) is a product that adjusts the microwave-infrared (IR) merged algorithm (hereafter GSMaP_MVK) with the global gauge analysis (CPC Unified Gauge-Based Analysis of Global Daily Precipitation) supplied by NOAA. The product also has a spatial and temporal resolution of 0.1 degree and 1 h. The 24-h accumulation product (12 UTC to 12 UTC) and version 7, released in 2017, is used for this study.

2.6. Statistical and Categorical Indexes

Statistical and categorical indexes [46] are used to evaluate the TMPA-V7 (version 7), IMERG-F (version 5) and GSMaP-G (version 7) products. Tables 1–3 give short descriptions of those statistic parameters, while an intensity rain classification based on daily thresholds is presented in Table 4.

Table 1. Statistical indexes.

Statistical Index	Equation	Optimal Value
Root Mean Square Error	$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(E_i - O_i)^2}{N}}$	0
Mean Error	$ME = \frac{1}{N}\sum_{i=1}^{N}(E_i - O_i)$	0

Table 2. Contingency Table.

	Gauge Rain	Gauge No-Rain	Total
Satellite rain	a = Hit	b = false alarm	E = (a + b)
Satellite no-rain	c = miss	d = correct negative	(c + d)
Total	O = (a + c)	(b + d)	(a + b + c + d)

Table 3. Categorical indexes.

Categorical Index	Equation	Optimal Value
Adjusted Equitable Threat Score (Mesinger, 2008)	$ETS = \frac{\left(H_a - \frac{O*O}{N}\right)}{\left(E+O+H_a-\frac{O*O}{N}\right)}$ Where, $H_a = O * \left[1 - \left(\frac{O-H}{O}\right)^{\frac{O}{E}}\right]$	1
Probability of detection	$POD = \frac{H}{O}$	1
False alarm ratio	$FAR = \frac{E-H}{E}$	0
BIAS	$BIAS = \frac{E}{O}$	1
Critical success index	$CSI = \frac{H}{E+O-H}$	1

Table 4. Rain classification and thresholds.

Rain Intensity Classification	Precipitation Thresholds (mm)
Rain/no-rain	0.5
Light	2–5
Moderate	10–20
Heavy	35–50

2.7. Standardization of Data

Data used in this validation study are generated in different formats and spatial resolutions. GSMaP-G, IMERG-F and TMPA-V7 products are regularly spaced, although with degree resolutions of 0.10 for the first two and 0.25 for the third. Observations (OBS) are measured in fixed points (latitude, longitude and precipitation value), not following a regular spatial pattern, which requires a standardization of the dataset. In this study, we chose to evaluate the products of precipitation estimates in a coarser resolution grid (0.25 degrees). Standardization was performed following these steps:

(a) Using the position (latitude and longitude) of each station, satellite-based precipitation retrievals are extracted from TMPA-V7, IMERG-F and GSMaP-G products using the nearest neighbor approach (the closest center of the correspondent grid point is selected). This approach is the same as that used in [11] to retain the original retrieved value of each algorithm. In this case, the maximum distance between the center of the grid point and the gauge is approximately seven kilometers for IMERG-F and GSMaP-G, and eighteen kilometers for TMPA-V7 (below the nominal spatial resolution of the respective products);

(b) A table is built with the latitude and longitude of the station, observed precipitation and estimated precipitation for the three products following the procedure described in the paragraph above;

(c) From this table, and using the same regularly spaced grid as that of the TMPA-V7 (0.25° × 0.25°), three grids with the averages of existing precipitations inside each grid point are calculated for IMERG-F, GSMaP-G and OBS. In the case of the TMPA-V7, the original value is preserved. These values represent the average precipitation at each grid point. Grid points with no existing gauges are flagged as invalid. Additionally, the average of the brightness temperature of GOES-13 channel 4 (10.8 microns) is also performed for those grid points with at least one gauge station. This variable, which represents the temperature of the top of the cloud, is used as a proxy to identify, in a very general way, the mean depth of the clouds;

(d) In order to perform a statistically robust study, only grid points with 50% or more of rain gauge data frequency, using the entire time series, were considered. The spatial distribution of points which satisfy this criterion is shown in Figure 2. Table 5 shows the amount of valid grid points per region.

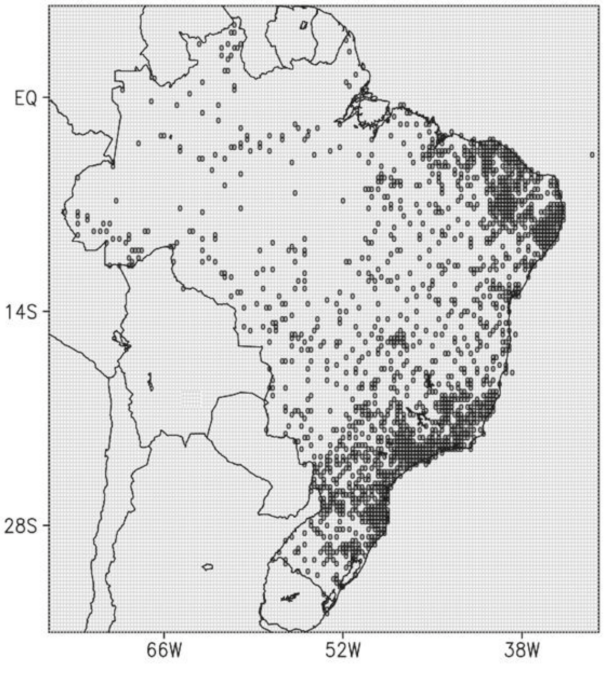


Figure 2. Spatial distribution of grid points which present rain gauge data frequency of at least 50% in the studied period, and in regularly spaced grids of 0.25°.

Table 5. Number of valid grid points for each region.

Region	N. of Grid Points
R1	271
R2	892
R3	270
R4	222
R5	124
BRAZIL	1779

3. Results

3.1. Temporal Evolution

Figure 3 shows the temporal evolution of the daily averages of TMPA-V7, IMERG-F, GSMaP-G observed precipitation (gauges), and brightness temperature from GOES-13 (Geostationary Operational Environmental Satellite) for Brazil and its five precipitation sub-regions. Brightness temperature is used here as a proxy in order to identify cloud top types (cold or warm) in each region. In order to smooth the higher frequencies of the time-series of each variable, a 10-day moving average was applied to all aforementioned variables.

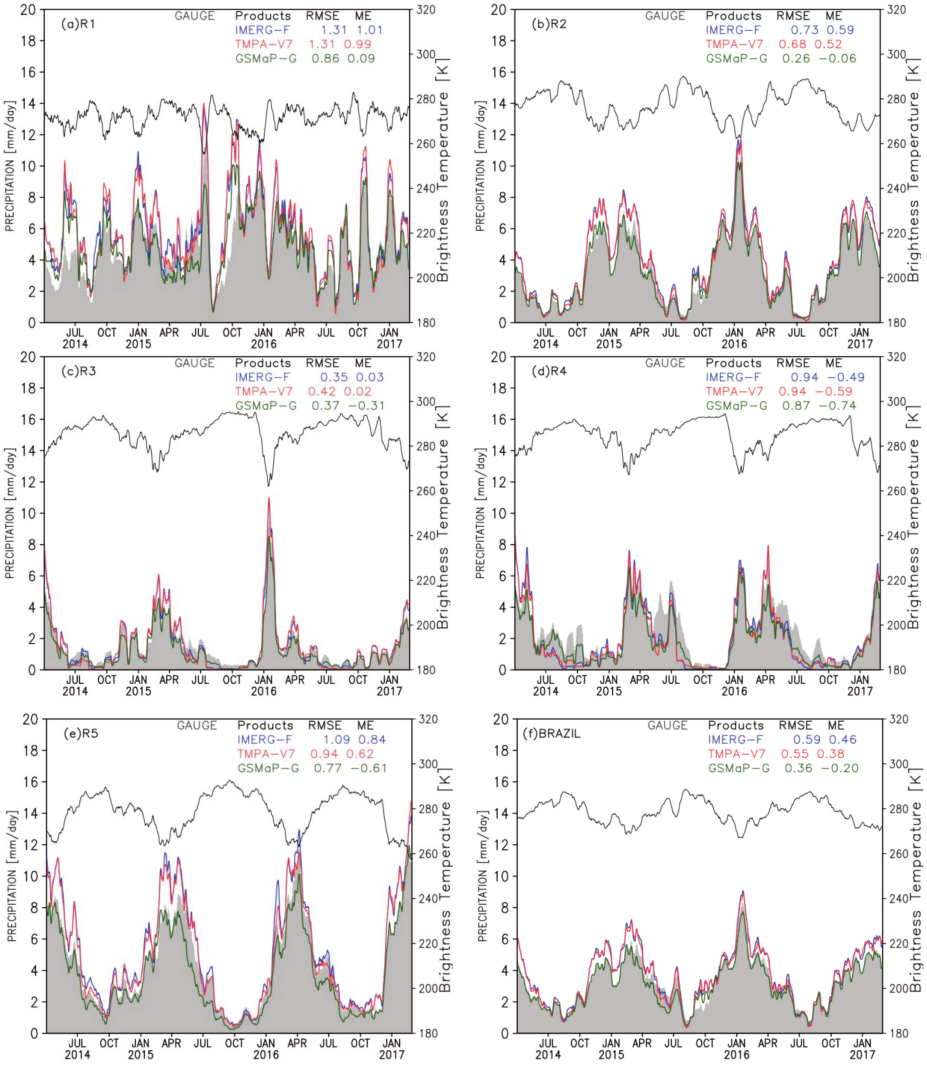


Figure 3. Temporal evolution of daily averages of the precipitation estimate products IMERG-F (blue), TMPA-V7 (red) and brightness temperature (black) from GOES-13 (Geostationary Operational Environmental Satellite) for the R1 (**a**), R2 (**b**), R3 (**c**), R4 (**d**), R5 (**e**) regions, and whole Brazil (**f**).

In R1 region (Figure 3a), considering the 271 valid grid points, TMPA-V7 and IMERG-F show similar behavior: both algorithms overestimate the accumulated daily precipitation. The overestimation is more evident from 2014 to the end of 2016, compared to that of the rest of the period. During this period, brightness temperature is highly correlated (negatively) with observed and estimated precipitation. The periods with larger overestimation of precipitation correspond to the lowest brightness temperature values. This behavior suggests the presence of deep convection systems (with large amount of ice on its structure), where larger amount of rainfall is estimated than is actually occurring. In this case, both products have the same RMSE value (1.31 mm/day), while the ME associated with TMPA-V7 (0.99 mm/day) is slightly lower than IMERG-F (1.01 mm/day). In the case of the GSMaP-G, the mean value is always below TMPA-V7 and IMERG-F retrievals, and close to observed values. This result suggests that the CPC Unified Gauge-Based Analysis of Global Daily precipitation used to adjust the bias in GSMaP-G (Ushio et al., 2013) is performing better than the monthly accumulation used to scale the NASA final products (TMPA-V7 and IMERG-F). For this region, ME is close to 0 and RMSE is 0.86 mm/day for GSMaP-G.

R2 region (Figure 3b), with 892 valid grid points, is the largest evaluated region. The behavior of maximum precipitation in summer and minimum in winter is estimated efficiently by all products. However, overestimation is also present along the studied period (mainly during summer time) for TMPA-V7 and IMERG-F, while GSMaP represents that peak better. As within R1 region, R2 region is influenced by deep convective systems, and the same behavior was seen: larger precipitation values were observed during periods with minimum brightness temperatures (generally summer months). In this region, RMSE (0.68 mm/day) and ME (0.52 mm/day) values for TMPA-V7 were slightly lower than RMSE (0.73 mm/day) and ME (0.59 mm/day) values from IMERG-F, while GSMaP-G has the lower values for all algorithms: 0.26 mm/day for RMSE and −0.06 mm/day for ME.

R3 region (Figure 3c), with 270 valid grid points, presents the lowest precipitation totals among all regions. All products showed similar behavior and are able to effectively estimate the precipitation regime in most of the analyzed period, except for a general tendency for overestimation during the wet season by TMPA-V7 and IMERG-F, and underestimation by GSMaP-G. However, in terms of ME, IMERG-F and TMPA-V7 have the lowest values, with 0.03 and 0.02 mm/day respectively. GSMaP-G has a negative bias of −0.31 mm/day, due to underestimation during the transition from the wet to the dry season (approximately April-August each year). RMSE values show that IMERG-F is slightly better that GSMaP-G (0.35 an 0.37 mm/day respectively) while TMPA-V7 has the worst performance (0.42 mm/day).

R4 (with 222 valid grid points) present the least accurate estimate of the precipitation regime of all products (Figure 3d). As observed in other regions, the behavior of all products is similar. During most of the period, there is an underestimation in precipitation values from all products, especially in days when brightness temperatures are higher (above 285 K). In this situation, precipitation is associated with systems with warm cloud tops. However, for precipitation systems characterized by cold cloud tops (relative minimum values of brightness temperature), better results are obtained. In this region, TMPA-V7 and IMERG-F show the same RMSE value (0.94 mm/day), but the ME is slightly lower for IMERG-F (−0.49 mm/day). For GSMaP-G, RMSE is 0.87 mm/day, the best value among all algorithms, whereas ME is −0.74 mm/day, the largest (negative) bias for all products.

R5 region (Figure 3e), located in the extreme north of the country, presents the lowest coverage of observed data (124 valid grid points). The annual precipitation cycle, with maximum values in March/April and minimum values in October, is relatively well represented, although with some overestimation for TMPA-V7 and IMERG-F and underestimation for GSMaP-G. In this particular region GSMaP-G has the best performance in terms of RMSE: 30% and 18% better than IMERG-F and TMPA-V7 respectively. In this region, RMSE (0.94 mm/day) and ME (0.62 mm/day) values for TMPA-V7 were slightly lower than RMSE (1.09 mm/day) and ME (0.84 mm/day) values from IMERG-F, while GSMaP-G has lower values for all algorithms: 0.77 mm/day for RMSE and −0.61 mm/day for

ME. A negative correlation between brightness temperature and the precipitation estimate is quite evident in this region.

Considering the whole Brazilian territory (Figure 3f), results are quite similar to those found in region R2, with maximum precipitation totals in January and minimum totals in July. This occurs due to the largest extent of region R2 in area and valid grid point of observations (about 59% of the total).

3.2. Quantitative Precipitation Forecast (QPF)

In this section, quantitative results will be presented for eight precipitation thresholds—0.5, 2, 5, 10, 15, 20, 35 and 50 mm, based on a contingency table (Table 3). As with the previous section, only grid points with at least 50% of observed data are considered in this analysis.

Figure 4 shows the Equitable Threat Score (ETS) for both evaluated products. ETS analysis for region R1 (Figure 4a) shows that GSMaP-G presented the best performance for all precipitation thresholds. The lowest ETS values, as expected, are observed during intense precipitation episodes (more than 50mm). In general, the ETS for GSMaP-G is about 20% higher than IMERG-F, and 35% better than TMPA-V7. Considering that ETS measures the fraction of observed grid points for a given threshold that were correctly estimated by a given algorithm, adjusted for hits associated with random chance, this score is showing that daily gauge adjustment used by GSMaP-G is adding some extra value compared with monthly-adjusted algorithms (TMPA-V7 and IMERG-F). Because ETS allows scores to be compared more fairly across different regimes, it is possible to conclude that R1 presents the best performance for all products.

R2 region (Figure 4b) presents ETS values slightly lower than in R1, showing that precipitation estimates in region R1 are more accurate. The GSMaP-G product exhibits better performance for all precipitation thresholds. Above 2 mm threshold, all products show a considerable decrease in performance. In most of the thresholds, the performance of GSMaP-G is about 30% better than IMERG-F, while TMPA-V7 is always on the low end for all thresholds.

Inner northeastern Brazil, represented by region R3 (Figure 4c), shows different behavior compared to that of regions R1 and R2. While GSMaP-G has an almost constant value up to 5 mm, an increase in ETS values from the 0.5 to the 5.0 mm threshold is observed for TMPA-V7 and IMERG-F. However, there is a decrease in performance, as expected. In all thresholds, GSMaP-G is superior to IMERG-F and TMPA-V7. This is more evident for lower rain rates (below 5 mm/day). The performance of all products in this region is considerably lower than that in regions R1 and R2.

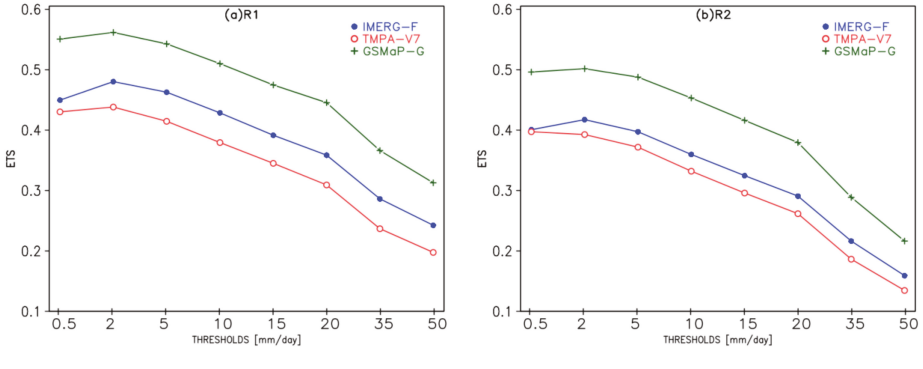


Figure 4. *Cont.*

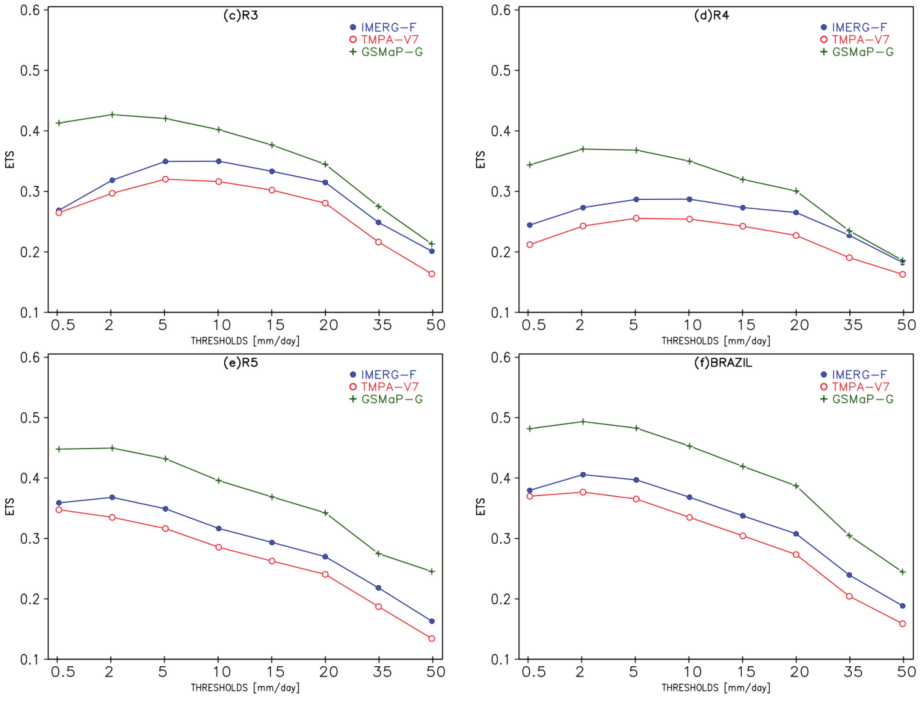


Figure 4. Equitable Threat Score (ETS), considering the whole studied period, for regions R1 (**a**), R2 (**b**), R3 (**c**), R4 (**d**), R5 (**e**) and Brazil (**f**).

The Eastern coast of northeastern Brazil, corresponding to region R4 (Figure 4d), shows the worst performance for all products among all regions. The performance of GSMaP-G is higher in this region for all thresholds, with a similar pattern to region R3. This might be attributed to the presence of warm clouds system. Palhiarini and Vila (2017) concluded, using 17 years of TRMM-PR (TRMM-precipitation radar) data, that shallow convection is the predominant cloud type system during the rainy season in this region. The lack of ice in precipitating clouds is this region makes it very difficult for microwave sensors to retrieve rainfall using high frequency channels (Braga, 2014).

In region R5 (Figure 4e), the ETS values for GSMaP-G are higher than those of NASA products. When considering the whole country (Figure 4f), the ETS values, as expected, are higher for GSMaP-G for all precipitation thresholds.

Figure 5 presents the performance diagram [47] of IMERG-F (blue), GSMaP-G (green) and TMPA-V7 (red) products. The circles represent the precipitation thresholds. The smallest circle represents the rain-no rain threshold, and the largest circle, the threshold above 50 mm. This diagram makes it easier to analyze the results, as it represents several dichotomic (yes/no) quality measures simultaneously, such as the POD, BIAS, CSI and FAR. Dashed lines represent BIAS, and solid lines the CSI. Thus, the best estimates are located in the upper right part of the diagram.

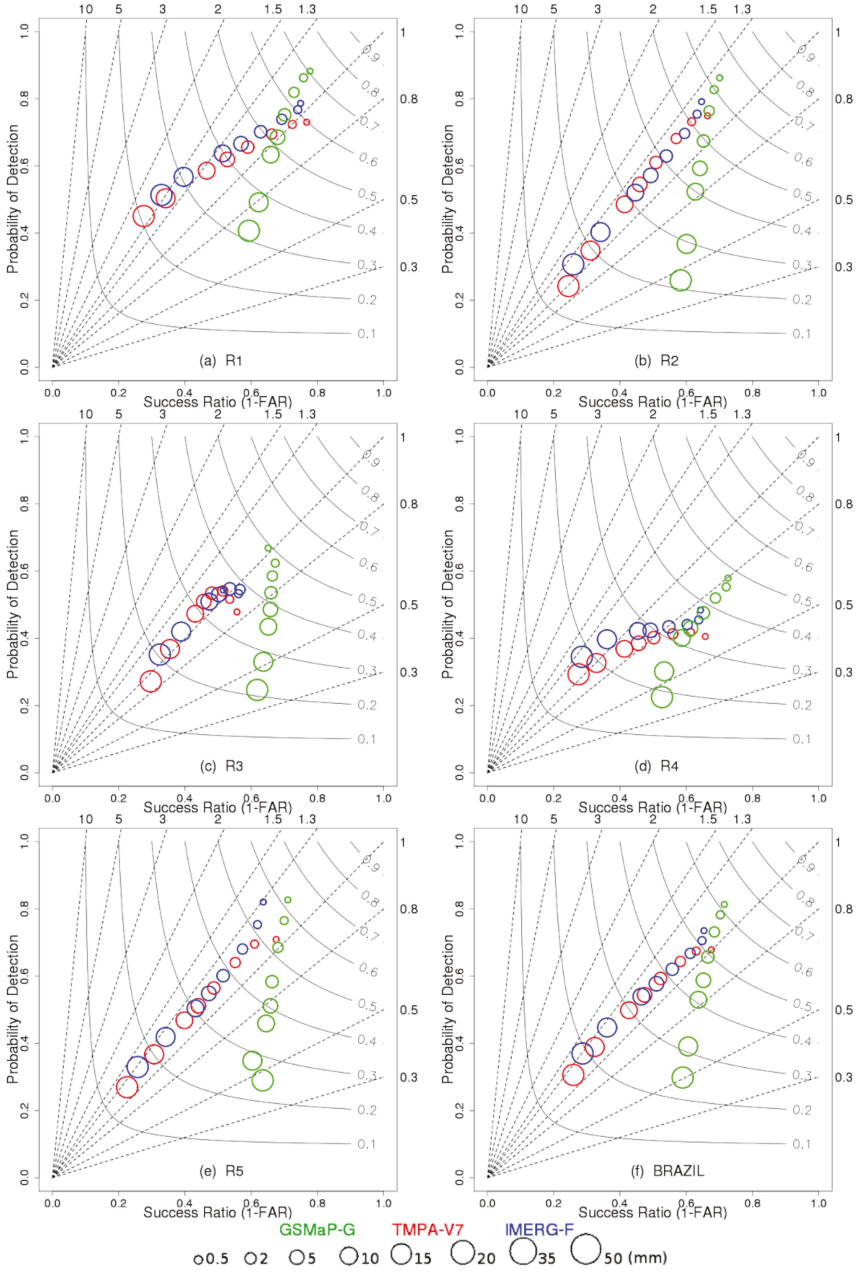


Figure 5. Performance diagram [47] summarizing the SR, POD, BIAS, and CSI for regions R1 (**a**), R2 (**b**), R3 (**c**), R4 (**d**), R5 (**e**) and BRAZIL (**f**). Dashed lines represent BIAS scores with labels on the outward extension of the line, while labelled solid contours are CSI. Circles represent the eight precipitation thresholds. The smallest circle represents the rain/no rain threshold (0.5 mm), and the largest circle represents the threshold above 50 mm.

The CSI for R1 region (Figure 5a) indicates better performance for the GSMaP-G product, followed by the GMP_F, and finally by TMPA-V7 for all thresholds analyzed. CSI for GSMaP-G is around 13% higher than the IMERG-F, and 20% higher than the TMPA-V7. In terms of FAR, it is observed that for thresholds below 2.0 mm, the results are similar among all three products, but for higher thresholds, FAR values for GSMaP-G are significantly lower. When this result is combined with a BIAS score below 0.8, indicates that GSMaP-G is missing high precipitation events in this region, while the relatively higher FAR value for NASA products, combined with BIAS around 1.5 is related with overestimation of the area with heavy precipitation rates (20–50 mm/day).

A similar analysis could be done for R2 (Figure 5b): GSMaP-G has better performance in terms of CSI, followed by IMERG-F and TMPA-V7. However, GSMaP-G is missing moderate and heavy rain events (low FAR and BIAS < 1), while IMERG-F and TMPA-V7 slightly overestimate the area for moderate and light rain rates.

For R3 region (Figure 5c), as was observed for R1 and R2, GSMaP-G exhibits higher CSI values for all thresholds when compared to the other two products. However, the underestimation of the area with rainfall (BIAS < 1) is present in GSMaP-G for all thresholds. The BIAS for IMERG-F and TMPA-V7 are close to 1, which means that the area with precipitation above a given threshold is correctly estimated. Low POD and high FAR values for high rain rates suggest that those high impact events are not correctly placed for NASA products.

Among all the analyzed regions, the R4 region (Figure 5d) shows the lowest performance. The same behavior observed in other regions is observed here. The main difference is the underestimation of the area for IMERG-F and TMPA-V7 for moderate and light rain rates.

R5 region (Figure 5e) has a similar patter than R2, the largest analyzed region. GSMaP-G has better performance in term of CSI, followed by IMERG-F and TMPA-V7. However, GSMaP-G is missing moderate and heavy rain events (low FAR and BIAS < 1), while IMERG-F and TMPA-V7 slightly overestimate the area for all rainfall threshold.

It is not surprising that Brazil, as a whole region (Figure 5d), shows a similar pattern to that of the R2 region. Because the largest number of gauges are found in R2, the weight in the final result is larger than other regions. However, the behavior of the analyzed categorical indexes is quite similar along all different rainfall regimes.

4. Discussion

The performance of precipitation estimate products, obtained using three state-of-the-art algorithms for GPM era, namely IMERG-F and TMPA-V7 from NASA and GSMaP-G from JAXA, were evaluated in five Brazilian regions with different precipitation regimes, and the whole Brazilian territory, during the period from April 2014 to February 2017. From a broad perspective, and considering the amount of rainfall, GSMaP-G has the lowest ME and RMSE when compared with NASA products for most of the regions, except for R4. This is also true for categorical indexes like CSI and ETS. However, when the area of rainfall above a certain threshold is evaluated, GSMaP tends to miss moderate and heavy precipitation events in almost all areas. These results are in a good agreement with other studies like [48], where IMERG-F overestimates extreme precipitation indices, but GSMaP-G shows a significant underestimation in several basins in China.

In this section, the possible reasons for this behavior will be discussed, considering the characteristics of each database. In this study, gauge bias-adjusted versions were selected as the final products for each algorithm to make a fair comparison. However, while for IMERG-F and TMPA-V7 this procedure is done using monthly totals from GPCC [13], in the case of GSMaP, a daily gauge analysis (CPC Unified Gauge-Based Analysis of Global Daily Precipitation) supplied by CPC/NOAA is used.

In the first case, the hourly accumulated rainfall (or 3-hourly, in case of TRMM-V7) is obtained using the monthly GPCC precipitation gauge analysis (over land) in a three-step process. First, the gauge analysis is adjusted by multiplying the monthly precipitation values with the corresponding

month's gridbox climatological adjustment ratios. Second, the multi-satellite estimate is adjusted to the large-scale mean of the gauges. Finally, the adjusted multi-satellite and gauge fields are combined using weighting by inverse estimated error variance [4]. In the case of GSMaP-G, the CPC Unified Gauge-Based Analysis of Global Daily Precipitation is applied based on the optimal theory which adjusts the GSMaP-Gauge hourly rain rate, so that the sum of the 24 h GSMaP-G rain rate is roughly same as the gauge measurement where those gauges are available [33].

The observational database used as ground truth, as described in Section 2.2, is composed of the global telecommunication system (GTS), the automatic platform for collecting data (PCDs), and the gauge data from regional meteorology centers in Brazil. Because NOAA receives some of those gauges in real time, they are also included in the CPC Unified Gauge-Based Analysis of Global Daily Precipitation product. In such cases, the observational and the GSMaP-G database are not completely independent. During the period of this study, the mean number of gauges used for validation was around 3400 per day, while CPC Unified Gauge-Based Analysis of Global Daily Precipitation uses approximately 1000 gauges per day (on average). This could explain the lower ME and RMSE of this algorithm compared to those of NASA products for the regions with largest amount of gauges (R1 and R2). However, this fact does not exclude this product from the comparison, because, even considering this limitation, this algorithm shows some characteristics which are interesting to consider: (i) it represents state-of-theart of satellite rainfall retrievals at JAXA (a partner of GPM program) with continuous developments and reprocessing cycles; (ii) it does not reproduce for all times, all regions and all periods, the evolution of the observed precipitation, which means that some degree of independent data remains in the database (Figure 3), and (iii) this gauge-adjusted product is the only one which is available in near real-time (4 days), compared with the two months' latency for IMERG-F and TMPA-V7. This last issue is quite important for some applications, such as hydrological forecasting (dams management) for hydropower and irrigation, and other users where the latency plays a vital role in the decision making process and for which the accuracy of retrieval should be above certain threshold which cannot be reached using the satellite-only versions (real time versions).

In the case of NASA products, the mean number of gauges available in the GPCC database used for bias correction in the final products over Brazil for the period 2014–2017 is around 300. This number is less than 10% of gauge data available for validation (~3400). Even those databases (TMPA-V7 and IMERG-F) are also not completely independent; a larger 'degree of independence' is achieved when compared with GSMaP-G.

Future validation studies should include no gauge-adjusted versions of these algorithms and fully independent observed data (i.e., radar estimates), to come to more conclusive results about the performance of these algorithms.

5. Conclusions

The performance of three satellite-based rainfall estimation products were evaluated in five Brazilian regions with different precipitation regimes, and the whole Brazilian territory, during the period from April 2014 to February 2017. Generally speaking, all products are able to estimate, with different degrees of accuracy, the levels of precipitation over the Brazilian territory. While overestimations are present in most of the studied regions for NASA products, GSMaP-G tend to slightly underestimate the observed rainfall. The most noticeable estimation errors for all products occur over the eastern coast of northeastern Brazil (region R4), where large underestimation for all products occurred during precipitation episodes caused by warm clouds. The quantitative analysis (ETS and CSI) shows that the GSMaP-G product presents better performance in all regions and all precipitation thresholds, while large underestimation of the area covered with heavy rainfall (rain rates > 100mm/day) is also observed for this algorithm (Figure 5). IMERG-F and TMPA-V7 show similar behavior in terms of CSI, ETS, POD, FAR and BIAS with a better performance for IMERG-F.

This study shows that GSMaP-G and IMERG-F precipitation products exhibit better performance compared to the current TMPA-V7, besides the finer horizontal and temporal resolution of the new

generation products. In this context, GSMaP-G and IMERG-F algorithms are a great replacements for TMPA-V7 products in the Brazilian territory, characterized by high density of river basins throughout its territory, where flood and landslide events are common, with negative social and environmental impacts. However, the choice of a given product will depend on the user's needs: GSMaP-G has a lower latency and tends to represent better to total amount of rainfall, while IMERG-F is more accurate for the retrieval of moderate and heavy rainfall events in terms of frequency (area).

Acknowledgments: The authors express their sincere thanks to the scientists in the National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA), responsible for the development of IMERG-F, TMPA-V7 and GSMaP-G algorithms and for providing satellite precipitation data. They also acknowledge the Center for Weather Forecast and Climate Studies for the gauge data database utilized in this study. The second author would like to acknowledge to Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) project 2015/14497-0 "Previsão Imediata de Tempestades Intensas e Entendimento dos Processos Físicos no Interior das Nuvens: O SOS-CHUVA (Sistema de Observação e Previsão de Tempo Severo)".

Author Contributions: José Roberto Rozante and Daniel A. Vila conceived the structure of this paper, and contributed to the discussion of scientific problems and the analysis of the results. Júlio Barboza Chiquetto, Alex de A. Fernandes and Débora Souza Alvim contributed to the data processing. All the authors conducted the manuscript revision.

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

References

1. Kelley, O.A. Where the Least Rainfall Occurs in the Sahara Desert, the TRMM Radar Reveals a Different Pattern of Rainfall Each Season. *J. Clim.* **2014**, *27*, 6919–6939. [CrossRef]
2. Knox, R.; Bisht, G.; Wang, J.; Bras, R. Precipitation Variability over the Forest-to-Nonforest Transition in Southwestern Amazonia. *J. Clim.* **2011**, *24*, 2368–2377. [CrossRef]
3. Short, D.A.; Nakamura, K. TRMM Radar Observations of Shallow Precipitation over the Tropical Oceans. *J. Clim.* **2000**, *13*, 4107–4124. [CrossRef]
4. 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]
5. 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]
6. 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]
7. Aonashi, K.; Awaka, J.; Hirose, M.; Kozu, T.; Kubota, T.; Liu, G.; Shige, S.; KIDA, S.; Seto, S.; Takahashi, N.; Takayabu, Y.N. GSMaP Passive Microwave Precipitation Retrieval Algorithm: Algorithm Description and Validation. *J. Meteorol. Soc. Jpn.* **2009**, *87A*, 119–136. [CrossRef]
8. Barrett, E.C. The estimation of monthly rainfall from satellite. *Mon. Weather Rev.* **1970**, *98*, 322–327. [CrossRef]
9. Follansbee, W.A. Estimation of Average Daily Rainfall from Satellite Cloud Photographs. 1973. Available online: https://trove.nla.gov.au/work/17767436?selectedversion=NBD6944042 (accessed on 21 August 2017).
10. Kummerow, C.; Barnes, W.; Kozum, T.; Shiue, J.; Simpson, J. The Tropical Rainfall Measuring Mission (TRMM) Sensor Package. *J. Atmos. Ocean. Technol.* **1998**, *15*, 809–817. [CrossRef]
11. 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: Dordrecht, The Netherlands, 2010; pp. 3–22.
12. 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. Version* **2015**, *4*, 30.
13. Schneider, U.; Becker, A.; Finger, P.; Meyer-Christoffer, A.; Ziese, M.; Rudolf, B. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.* **2014**, *115*, 15–40. [CrossRef]
14. Hamada, A.; Murayama, Y.; Takayabu, Y.N. Regional Characteristics of Extreme Rainfall Extracted from TRMM PR Measurements. *J. Clim.* **2014**, *27*, 8151–8169. [CrossRef]

15. Agel, L.; Barlow, M.; Qian, J.-H.; Colby, F.; Douglas, E.; Eichler, T. Climatology of Daily Precipitation and Extreme Precipitation Events in the Northeast United States. *J. Hydrometeorol.* **2015**, *16*, 2537–2557. [CrossRef]
16. Huang, Y.; Chen, S.; Cao, Q.; Hong, Y.; Wu, B.; Huang, M.; Qiao, L.; Zhang, Z.; Li, Z.; Li, W.; et al. Evaluation of Version-7 TRMM Multi-Satellite Precipitation Analysis Product during the Beijing Extreme Heavy Rainfall Event of 21 July 2012. *Water* **2013**, *6*, 32–44. [CrossRef]
17. Li, L.; Hong, Y.; Wang, J.; Adler, R.F.; Policelli, F.S.; Habib, S.; Irwn, D.; Korme, T.; Okello, L. Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Nat. Hazards* **2009**, *50*, 109–123. [CrossRef]
18. Dinis, P.A.; Mantas, V.; Andrade, P.S.; Tonecas, J.; Kapula, E.; Pereira, A.; Carvalho, F.S. Contribution of TRMM rainfall data to the study of natural systems and risk assessment. Cases of application in SW Angola. *Estudos Quaternário* **2013**, *9*, 33–43.
19. Siddique-E-Akbor, A.H.M.; Hossain, F.; Sikder, S.; Shum, C.K.; Tseng, S.; Yi, Y.; Turk, F.J.; Limaye, A. Satellite Precipitation Data–Driven Hydrological Modeling for Water Resources Management in the Ganges, Brahmaputra, and Meghna Basins. *Earth Interact.* **2014**, *18*, 1–25. [CrossRef]
20. Rozante, J.R.; Cavalcanti, I.F.A. Regional Eta model experiments: SALLJEX and MCS development. *J. Geophys. Res.* **2008**, *113*, D17106. [CrossRef]
21. Da Rocha, R.P.; Morales, C.A.; Cuadra, S.V.; Ambrizzi, T. Precipitation diurnal cycle and summer climatology assessment over South America: An evaluation of Regional Climate Model version 3 simulations. *J. Geophys. Res.* **2009**, *114*, D10108. [CrossRef]
22. Nicholson, S.E.; Some, B.; McCollum, J.; Nelkin, E.; Klotter, D.; Berte, Y.; Diallo, B.M.; Gaye, I.; Kpabeba, G.; Ndiaye, O.; et al. Validation of TRMM and Other Rainfall Estimates with a High-Density Gauge Dataset for West Africa. Part II: Validation of TRMM Rainfall Products. *J. Appl. Meteorol.* **2003**, *42*, 1355–1368. [CrossRef]
23. Zou, C.-Z.; Zheng, W. Simulation of diurnal patterns of summer precipitation in the North American monsoon: An assessment using TRMM. *Geophys. Res. Lett.* **2004**, *31*. [CrossRef]
24. Wolff, D.B.; Marks, D.A.; Amitai, E.; Silberstein, D.S.; Fisher, B.L.; Tokay, A.; Wang, J.; Pippitt, J.L. Ground Validation for the Tropical Rainfall Measuring Mission (TRMM). *J. Atmos. Ocean. Technol.* **2005**, *22*, 365–380. [CrossRef]
25. Dinku, T.; Ceccato, P.; Grover-Kopec, E.; Lemma, M.; Connor, S.J.; Ropelewski, C.F. Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* **2007**, *28*, 1503–1526. [CrossRef]
26. Su, F.; Hong, Y.; Lettenmaier, D.P. Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin. *J. Hydrometeorol.* **2008**, *9*, 622–640. [CrossRef]
27. Franchito, S.H.; Rao, V.B.; Vasques, A.C.; Santo, C.M.E.; Conforte, J.C. Validation of TRMM precipitation radar monthly rainfall estimates over Brazil. *J. Geophys. Res.* **2009**, *114*, D02105. [CrossRef]
28. Palharini, A.; Santos, R.; Vila, D.A. Climatological Behavior of Precipitating Clouds in the Northeast Region of Brazil. *Adv. Meteorol.* **2017**, *2017*, 12. [CrossRef]
29. Laing, A.G.; Michael, F.J. The global population of mesoscale convective complexes. *Q. J. R. Meteorol. Soc.* **1997**, *123*, 389–405. [CrossRef]
30. 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]
31. Hsu, K.; Gao, X.; Sorooshian, S.; Gupta, H.V. recipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteorol.* **1997**, *36*, 1176–1190. [CrossRef]
32. 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]
33. Mega, T.; Ushio, T.; Kubota, T.; Kachi, M.; Aonashi, K.; Shige, S. Gauge Adjusted Global Satellite Mapping of Precipitation (GSMaP_Gauge). In *XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*; IEEE: Beijing, China, 2014; pp. 1–4.
34. Rozante, J.R.; Moreira, D.S.; de Goncalves, L.G.G.; Vila, D.A. Combining TRMM and Surface Observations of Precipitation: Technique and Validation over South America. *Weather Forecast.* **2010**, *25*, 885–894. [CrossRef]
35. Reboita, M.S.; Gan, M.A.; da Rocha, R.P.; Ambrizzi, T. Regimes de precipitação na América do Sul: Uma revisão bibliográfica. *Rev. Bras. Meteorol.* **2010**, *25*, 185–204. [CrossRef]

36. Velasco, I.; Fritsch, J.M. Mesoscale convective complexes in the Americas. *J. Geophys. Res.* **1987**, *92*, 9591. [CrossRef]
37. Quadro, M.F.L.; de Dias, M.A.F.; Herdies, D.L.; Gonçalves, L.G.G. de Análise climatológica da precipitação e do transporte de umidade na região da ZCAS através da nova geração de reanálises. *Rev. Bras. Meteorol.* **2012**, *27*, 152–162. [CrossRef]
38. Vera, C.; Baez, J.; Douglas, M.; Emmanuel, C.B.; Marengo, J.; Meitin, J.; Nicolini, M.; Nogues-Paegle, J.; Paegle, J.; Penalba, O.; et al. The South American Low-Level Jet Experiment. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 63–78. [CrossRef]
39. Zhou, J.; Lau, K.-M. Does a Monsoon Climate Exist over South America? *J. Clim.* **1998**, *11*, 1020–1040. [CrossRef]
40. Kousky, V.E. Alonso Gan M Upper tropospheric cyclonic vortices in the tropical South Atlantic. *Tellus* **1981**, *33*, 538–551. [CrossRef]
41. Kousky, V.E. Pentad outgoing longwave radiation climatology for the South American sector. *Rev. Bras. Meteorol.* **1988**, *3*, 217–231.
42. Cohen, J.C.P.; Silva Dias, M.A.F.; Nobre, C.A. Environmental Conditions Associated with Amazonian Squall Lines: A Case Study. *Mon. Weather Rev.* **1995**, *123*, 3163–3174. [CrossRef]
43. 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 Sens.* **2009**, *47*, 3084–3097. [CrossRef]
44. Kachi, M.; Kubota, T.; Ushio, T.; Shige, S.; Kida, S.; Aonashi, K.; Okamoto, K.; Oki, R. Development and Utilization of "JAXA Global Rainfall Watch" System based on Combined Microwave and Infrared Radiometers Aboard Satellites. *IEEJ Trans. Fundam. Mater.* **2011**, *131*, 729–737. [CrossRef]
45. Ushio, T.; Sasashige, K.; Kubote, T.; Shige, S.; Okamoto, K.; Aonashi, K.; Inoue, T.; Tankahashi, 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.* **2009**, *87A*, 137–151. [CrossRef]
46. Wilks, D. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Academic Press: Cambridge, MA, USA, 2011; p. 676.
47. Roebber, P.J. Visualizing Multiple Measures of Forecast Quality. *Weather Forecast.* **2009**, *24*, 601–608. [CrossRef]
48. Ning, S.; Song, F.; Udmale, P.; Jin, J.; Thapa, B.R.; Ishidaira, H. Error Analysis and Evaluation of the Latest GSMap and IMERG Precipitation Products over Eastern China. *Adv. Meteorol.* **2017**, *2017*, 1–16. [CrossRef]

remote sensing

MDPI

Article

SLALOM: An All-Surface Snow Water Path Retrieval Algorithm for the GPM Microwave Imager

Jean-François Rysman [1,*], Giulia Panegrossi [1], Paolo Sanò [1], Anna Cinzia Marra [1], Stefano Dietrich [1], Lisa Milani [2] and Mark S. Kulie [3]

[1] Institute of Atmospheric Sciences and Climate (ISAC)—National Research Council (CNR), 00133 Rome, Italy; giulia.panegrossi@artov.isac.cnr.it (G.P.); paolo.sano@artov.isac.cnr.it (P.S.); a.marra@isac.cnr.it (A.C.M.); s.dietrich@isac.cnr.it (S.D.)

[2] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; lisa.milani@nasa.gov

[3] Department of Geological and Mining Engineering and Sciences, Michigan Technological University, Houghton, MI 49931, USA; mkulie@mtu.edu

* Correspondence: jeanfrancois.rysman@artov.isac.cnr.it; Tel.: +39-06-4993-4492

Received: 19 June 2018; Accepted: 12 August 2018; Published: 14 August 2018

Abstract: This paper describes a new algorithm that is able to detect snowfall and retrieve the associated snow water path (SWP), for any surface type, using the Global Precipitation Measurement (GPM) Microwave Imager (GMI). The algorithm is tuned and evaluated against coincident observations of the Cloud Profiling Radar (CPR) onboard CloudSat. It is composed of three modules for (i) snowfall detection, (ii) supercooled droplet detection and (iii) SWP retrieval. This algorithm takes into account environmental conditions to retrieve SWP and does not rely on any surface classification scheme. The snowfall detection module is able to detect 83% of snowfall events including light SWP (down to 1×10^{-3} kg·m^{-2}) with a false alarm ratio of 0.12. The supercooled detection module detects 97% of events, with a false alarm ratio of 0.05. The SWP estimates show a relative bias of −11%, a correlation of 0.84 and a root mean square error of 0.04 kg·m^{-2}. Several applications of the algorithm are highlighted: Three case studies of snowfall events are investigated, and a 2-year high resolution 70°S–70°N snowfall occurrence distribution is presented. These results illustrate the high potential of this algorithm for snowfall detection and SWP retrieval using GMI.

Keywords: snowfall detection; snow water path retrieval; supercooled droplets detection; GPM Microwave Imager

1. Introduction

Snowfall plays a central role in the Earth's climate system. Indeed, it is directly connected to the occurrence of snow on the ground, which itself impacts the climate. First, the surface snow layer strongly affects the water cycle at high latitudes and in mountainous regions by storing water during the cold season and then releasing it progressively in the environment when the weather gets warmer. Secondly, snow covered areas impact the Earth's radiative budget. Indeed, snow covered surfaces have a much higher albedo than bare land and thus reflect a significant portion of incoming sun radiation, which limits the surface warming. Climate change could alter the spatial and temporal distribution of global snowfall patterns and snow cover. Even a slight increase in temperature can, in some cases, change the precipitation phase from snow to rain. In fact, several studies in the USA already reported a decrease of the proportion of snow events versus rain events in winter time [1,2]. This problem illustrates the strong need to better characterize snow distribution and variability at the global scale.

Global snowfall can only be monitored observationally using spaceborne instruments [3,4]. Spaceborne microwave sensors are particularly suitable for detecting and quantifying snowfall thanks to their unique ability to probe within clouds [5,6]. Active microwave sensors, such as the

Cloud Profiling Radar (CPR) onboard CloudSat [7] and the Dual-Frequency Precipitation Radar (DPR) onboard the Global Precipitation Measurement Mission-Core Observatory (GPM-CO, [8]), are particularly valuable for providing detailed vertical profiles of snow [9–13]. However, due to several limitations, such as their reduced swath (especially for CPR) or their limited sensitivity (for DPR), they are not well-suited to comprehensively characterizing snowfall. Specifically, the CPR swath width of 1.7 km implies a revisit time of about two weeks for a 100 km × 100 km pixel. In addition, CPR does not provide snowfall observations close to the surface (i.e., below ~1 km) due to ground clutter, which remains a substantial limitation for quantifying surface snowfall. Regarding DPR, Casella et al. [14] showed that this radar detects only 5–7% of global snowfall events with respect to CPR, while 29–34% of the CPR global snowfall mass is detected by DPR (for version 4 products). They also showed that, while DPR is mostly suitable for retrieving intense/deeper snowfall, and by optimally combining the dual-frequency signal (Ku and Ka band), DPR snowfall detection efficacy can increase significantly (up to 54–59% of the CPR snowfall mass), it misses a significant part of light snowfall events.

On the other hand, passive microwave sensors, such as the most recent GPM Microwave Imager (GMI) or the Advanced Technology Microwave Sounder (ATMS) (on board Suomi-NPP and NOAA-20 satellites), appear promising for snowfall characterization. These sensors have high frequency channels that are highly sensitive to snowfall due to the scattering by snowflakes of upwelling terrestrial and radiation, originating in the lower levels of the atmosphere [15–18]. In addition, passive microwave radiometers have a large swath and have been installed on many platforms over the last decades, ensuring a good global coverage with a fair spatial resolution and lengthy data records.

Therefore, several approaches have been proposed in recent years for detecting and retrieving snowfall using passive microwave radiometers, such as the Special Sensor Microwave-Humidity (SSM/T2) [19], Advanced Microwave Sounding Unit-B (AMSU-B) and Microwave Humidity Sounding (MHS) [5,16,20–22], Advanced Technology Microwave Sounder (ATMS) [23,24], Special Sensor Microwave-Imager/Sounder (SSMIS) [25] and GMI [26] (Figure 1). In addition, to these passive microwave products, specifically dedicated to snowfall, some operational products provide estimations of precipitation, including snowfall at the global scale, combining observations of several radiometers. In particular, the official NASA Goddard PROFiling (GPROF) Bayesian algorithm [27] retrieves precipitation for all the radiometers of the GPM constellation. GPROF is tuned using an a priori database of matching TBs and observed precipitation rates. These observed precipitation rates originate in DPR over vegetated lands and oceans, as well as in the Multi-Radar/Multi-Sensor (MRMS) U.S. ground radar network over snow covered surfaces. The phase classification of the precipitation (solid or liquid) is based on the parameterization scheme of Sims and Liu [28]. GPROF, as well as any algorithm based on the use of empirical datasets, suffers from the limitations of the products used as references. For example, GPROF snowfall retrievals are affected by the low sensitivity of DPR to snowfall (as shown in Casella et al. [14]), and by the difficulties of MRMS in retrieving surface snowfall and representing snowfall events at the global scale.

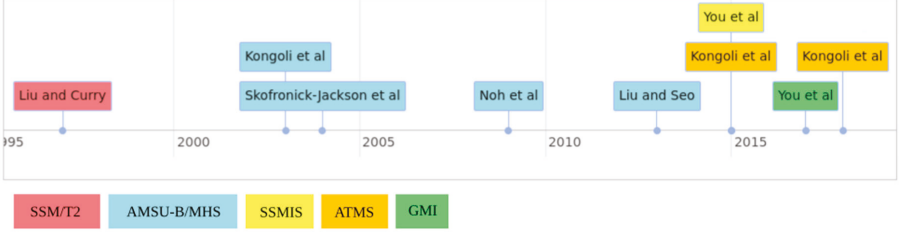


Figure 1. Timeline of several snowfall detection and retrieval algorithms. Box color indicates the instrument for which a given algorithm has been developed.

Despite these attempts, detecting and quantifying surface snowfall rates using spaceborne microwave radiometers remains a challenging task [3,26,29]. The high-frequency radiometer channel observations are very sensitive to environmental conditions (e.g., humidity, temperature, frozen or snow-covered soils), which affect the measured signal. This problem is particularly acute at high latitudes, where the low and variable emissivity of snow or ice-covered surfaces [22,30–32] can mask snowflakes' scattering signatures [33]. Moreover, low humidity in high latitude regions makes the atmosphere more transparent for channels that probe around the water vapor absorption line and thus increase surface contamination. The snow microphysics is also very complex, and a snow cloud is often composed of a gamut of snow particles with a variety of densities, shapes, particle size distributions and radiative properties [34–39]. In addition, supercooled droplets and melting snow frequently occur and can strongly affect the observed signal [16,40–42]. Finally, while the active sensors, such as CPR, provide vertical information on snow clouds, passive instruments only offer integrated information that combines the radiative signal of every cloud layer.

Recently, Panegrossi et al. [43] highlighted the impact of falling snow on the high frequency GMI channels using CloudSat CPR as a reference. This study focuses on higher latitude snowfall systems (around 60°N/S) and includes very weak snowfall events that are not considered in studies where DPR is used as a reference [26,33]. In particular, Panegrossi et al. [43] identified the environmental conditions under which snow detection could be optimally achieved (e.g., high atmospheric moisture and sea ice coverage) and characterized the sensitivity of GMI high-frequency channels and the polarization signal at 166 GHz to falling snow. They also highlighted the impact of supercooled droplets on the GMI signal, showing that a supercooled droplet layer on top of a snow cloud can filter out the scattering signal of the snow particles below. This study, as well as that of You et al. [26] and Ebtehaj et al. [33], have evidenced the need to characterize the extremely variable background surface for each GMI pixel at the time of the overpass, especially at high latitudes (in cold and dry conditions), where high-frequency channels may be affected by the emission and polarization signal from the surface. Interestingly, Panegrossi et al. [44] have shown that GMI low frequency channels can be useful for providing information about the background surface, such as the presence of sea ice and snow cover.

These results suggest that characterizing environmental conditions is essential to appropriately retrieving snowfall using passive microwave observations. In this context, we built a new algorithm, named SLALOM (for Snow retrievaL ALgorithm fOr gMi), to detect surface snowfall and retrieve the associated Snow Water Path (SWP) using GMI radiometer observations. The SLALOM algorithm is composed of three modules: The first, for surface snowfall detection, the second, for supercooled droplet detection and the third, for SWP retrieval. It is trained and validated against CloudSat CPR observations for surface snowfall detection and SWP retrieval, and against CloudSat CPR radar and CALIPSO CALIOP lidar for supercooled droplet detection. Section 2 is devoted to the description of the databases used and the methodology. The performance and characteristics of the algorithm are assessed in Section 3. Applications of the algorithm are shown in Section 4. The last section presents the conclusion and discussion.

2. Materials and Methods

The SLALOM algorithm is trained and evaluated using a coincident database between the GMI-GPM radiometer and the CPR-CloudSat radar. Several products, detailed in the following sections, are included in this database.

2.1. GMI-CPR Database

GMI is a conical scanning passive microwave radiometer that probes the atmosphere using 13 channels from 10.65 to 183.31 ± 7 GHz [45]. The 10.65, 18.7, 36.5, 89, and 166 GHz channels are equipped with vertical and horizontal polarizations. The field of view of the channels ranges from 19 km × 32 km (10.65 GHz) to 4.4 km × 7.2 km (channels from 89 to 183.31 ± 7 GHz) for a swath of about 900 km. This instrument flies onboard the GPM-Core Observatory (GPM-CO) satellite, launched

on 28 February 2014 [46]. This satellite orbits at an altitude of about 400 km, with a drifting orbit, for an inclination of 65°.

CPR is a 94.05 GHz cloud radar with a horizontal resolution of 1.4 (cross-track) km × 1.7 km (along track) and a vertical resolution of 480 m (oversampled to 240 m). This instrument was launched in April 2006 and has functioned only during daytime since April 2011. CPR is installed on-board the CloudSat satellite [7]. This satellite was part of the A-train constellation until February 2018 and flies with a near polar sun-synchronous orbit at an altitude of 700 km.

While GMI provides integrated information on the surface and atmosphere characteristics on a large swath, CPR furnishes observations on the whole atmospheric column, but with a very narrow swath (revisiting time of 16 days for a square of 100 km × 100 km). A coincident GMI-CPR dataset has recently been developed (2B-CSATGPM product, [47]) to leverage both instruments' respective strengths. For each ±15 min coincidence, CPR and DPR reflectivity profiles and GMI brightness temperatures (TBs) are archived.

The majority of the coincidences are found at around 60°N/S (Figure 2 (see also Figure 1 from Panegrossi et al. [43]). Numerous ancillary products are also available, including surface information (ground elevation, land-sea flag), CPR precipitation products [e.g., 2C-SNOW-PROFILE (2CSP)] and environmental variables [total precipitable water (TPW), 2-meter temperature (T2m)]. It is important to note that the 2CSP product, used in this study, retrieves snow water content (SWC), where snowfall at the surface is probable or certain, and, if estimated, the liquid fraction is <15% (dry snow). These conditions are based on interpolated European Centre for Medium-Range Weather Forecasts (ECMWF) model temperature profiles at the near-surface clutter-free bin and, as evidenced by Casella et al. [14], are therefore subject to the model uncertainty to the ground clutter conditions. Moreover, the 2CSP SWC is available only at or above the first clutter-free bin and only if the equivalent radar reflectivity factor at the first clutter-free bin is above −15 dBZ. In the following analyses, a snowfall event corresponds to a 2CSP near-surface snowfall rate of >0 mm/h. The coincident CPR/GMI dataset provides the TBs of every channel at their own resolution on the GMI S1 swath grid [with a distance between adjacent pixels of about 6 km (across track) × 13 km (along track)] [47]. In order to associate one single CPR profile with each GMI pixel, we identified the CPR profiles closest to each S1 pixels and averaged them. Finally, we enriched the GPM-CPR dataset with complementary measurements, which are detailed in the following section.

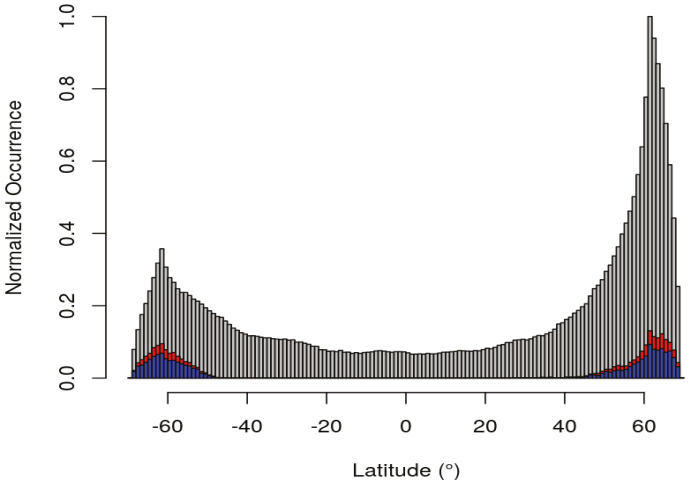


Figure 2. Zonal distribution of GMI/CPR coincidences: All (grey), snowfall (red), and snowfall with supercooled droplets (blue).

2.2. Complementary Dataset

Several studies (e.g., [40–43,48]) indicated that the presence of supercooled droplets on the top of the snow clouds could strongly affect TBs from high frequency channels. Specifically, supercooled droplets tend to partially mask scattering by snow crystals, which complicates the detection of snow using passive microwave radiometers. In order to take this into account, we used the water phase mask provided by the DARDAR (liDAR + raDAR) product. This product combines the observations of the CPR radar and of the CALIOP lidar (on-board CALIPSO) to provide insights, not only into the water phase, but also into the ice water content and ice particle effective radius, with a vertical (horizontal) resolution of 60 m [1.4 km (cross-track) $\times$ 1.7 km (along track)]. Thus, for each snow measurement in the GMI-CPR coincident database, we associated a flag for cases without supercooled droplets, cases with supercooled droplets embedded in the snow cloud and cases with supercooled droplets on top of the snow cloud. It should be noted that Battaglia et Delanoë [49] suggested that some supercooled layer occurrences might be missed by DARDAR due to lidar attenuation problems. Supercooled droplets that are embedded in or on top of the snow cloud are widespread at all latitudes, representing on average about 2/3 of snowfall events, as shown in Figure 2. Therefore, the presence of supercooled water and its vertical distribution within the cloud should be considered in snowfall retrieval algorithms.

Since the occurrence and content of snow are also strongly influenced by the vertical variability of environmental conditions, we also extracted the closest profiles of temperature, specific humidity and relative humidity from the European Reanalysis–Interim (ERAI) [50] at each point of the GMI-CPR dataset using a nearest-neighbour approach. The ERA-Interim dataset provides meteorological variable profile and surface estimations from 1979 to present, with a spatial resolution of $0.75^\circ \times 0.75^\circ$ and a temporal resolution of 6 hours. In order to reduce the number of variables for the vertical profiles, we computed the principal component of each of the 3 variables and retained the 4 first components, which represent more than 99% of the total variability.

Finally, Panegrossi et al. [43] showed that the sensitivity of TBs to snow depends on the characteristics of the background surface (e.g., open sea, sea ice and snow-cover). In order to evaluate the skill of the SLALOM algorithm for these different surface types, we added daily sea ice concentration observations derived from Advanced Microwave Scanning Radiometer (AMSR-E/AMSR2) measurements [51] to the GPM-CPR database.

3. SLALOM Algorithm

The SLALOM algorithm is composed of three modules for (i) snowfall detection, (ii) supercooled droplet detection, and (iii) SWP estimation. Snowfall and supercooled detection modules rely mainly on the random forest approach [52], while the SWP retrieval module uses a segmented multi-linear regression approach [53].

The random forest is a classification method that combines the outputs of numerous decision (regression) trees [54] to provide an optimal prediction of a given factor as a function of several input variables. A decision tree is itself a classification method that hierarchically splits the dataset into groups in order to minimize the variance within each group. Specifically, the decision tree chooses recursively one of the input variables that splits the dataset into 2 groups for which the variance is minimal (see also Section 5.1 for details). In general, a decision tree alone shows poor predictive skills. However, when several decision trees are combined, their combined predictive skill increases drastically. These qualities describe the general framework of the random forest algorithm. More specifically, the random forest algorithm builds several decision trees using random samples of the training dataset and of the input variables. Then, it combines the outputs of each tree to make an overall prediction. If all trees agree for the same prediction, the random forest algorithm provides a global prediction, with a probability of 100%. If trees disagree, the probability of the prediction decreases.

The segmented multi-linear regression method aims to find linear relationships between a response variable and several predictors. Regression is performed to minimize the difference

between the observations and the fitted linear relationship. This type of regression is called segmented (or piecewise) because different multi-linear relationships are chosen for different segments (range of values) of the predictors.

3.1. Snowfall Detection

For the snow detection module, we used the random forest algorithm to predict snowfall occurrence. The training dataset was comprised of 408,254 observations, with 38,331 2CSP-defined snowfall events. In order to build an optimal random forest model, we sought the best combinations of input variables, associated with the 13 GMI radiometer channels, in order to obtain the most reliable predictions. The input variables tested were sea ice concentration, T2m, TPW, the 4 first principal components of the vertical profiles of temperature, specific humidity and relative humidity, snow cover, and land-sea flag. Specifically, we selected 2000 random combinations of input variables and discarded those which were systematically associated with the lowest predictive scores. The optimal random forest chosen was built with the following input variables: T2m, TPW, the 4 first principal components of the vertical profiles of temperature, specific humidity and relative humidity, together with the 13 GMI radiometer channels. It should be noted that surface type variables were systematically associated with the lowest predictive scores in the selection process. This reveals that, by exploiting all GMI channels at the time of the overpass, the random forest itself can combine the information related to the surface conditions (deriving it from the low frequency channels) and the atmospheric variables (e.g., T2m and TPW) with the cloud vertical structure (i.e., presence or absence of snowfall). It should also be noted that, while 2CSP uses ERA-Interim temperature profiles to determine the phase of the detected precipitation, the SLALOM snowfall detection module uses the ERA-Interim temperature profile (together with the variables already listed) to detect snowfall occurrence. We evaluated the impact of the decision tree number of which the random forest (up to 1000 decision trees) is composed on the prediction skill. A 300 decision trees limit was adopted, since the predictive skill does not improve significantly over this number. The output of the snowfall detection module gives a probability of snowfall occurrence at the GMI high-frequency channel resolution, which is then used as an input for the second module (applied only to snowfall pixels) to detect supercooled droplets.

3.2. Supercooled Droplets Detection

Supercooled droplets detection is also based on a random forest approach. Note that, as opposed to Panegrossi et al. [43], we did not consider embedded supercooled droplet cases in the training dataset. Indeed, Panegrossi et al. [43] showed that the strongest impact of supercooled droplets on the measured TBs in the presence of snowfall is found when the supercooled droplets occur at the top of snow clouds. Conversely, the impact of supercooled droplets embedded in the cloud is reduced. The training dataset is the same as that for the snowfall detection module, except that it excludes any embedded supercooled droplet cases. Specifically, it contains 397,033 observations: 10,783 snowfall cases without supercooled droplets and 16,327 snowfall cases with supercooled droplets on top of the cloud. The optimal random forest is similar to the snow detection module, with the same input variables, i.e., T2m, TPW, the vertical profiles of temperature, specific humidity and relative humidity, all 13 GMI channels and 300 decision trees. The output of the supercooled droplet detection module gives a probability of supercooled droplet occurrence at the GMI high-frequency channel resolution, which is then used as an input for the SWP retrieval module.

3.3. SWP Retrieval

In this last module, we sought to infer SWP associated with snowfall using GMI high frequency channel observations. As GMI high frequency channels provide vertically integrated measurements of the atmosphere, the signal of TBs in the presence of snowfall is related to SWP. However, the relationship between SWP and GMI TBs is strongly affected by environmental conditions. Thus, it appeared meaningful to develop an SWP retrieval model that adapts to these conditions.

With this aim, we first used a decision tree to split the observations into subsets using the following input variables: T2m, TPW, supercooled occurrence flag and the low frequency GMI channels (i.e., 10.65 H&V, 18.7 H&V, 23.8 V, and 36.5 H&V). We ensured that each subset contained at least 300 observations in order to keep a significant number of points in every subset for the regression procedure. Then, for each of these subsets, we performed a segmented multi-linear regression of SWP as a function of high frequency channel TBs (89 H&V, 166 H&V, 183.31 $\pm$ 7, and 183.31 $\pm$ 3). Thus, this approach provides a different model for each of the defined subsets, each one of them being adapted to a specific environmental condition.

4. Algorithm Evaluation

We evaluated the predictive capabilities of the algorithm using an independent dataset (i.e., distinct from the training dataset), populated with a random selection of 20% from the observations of the initial dataset (9839 snowfall cases and 102,064 no-snowfall cases). The snowfall and supercooled detection are evaluated using the Probability of Detection (POD), False Alarm Ratio (FAR) and Heidke Skill Score (HSS) metrics [55]:

$$\text{Heidke Skill Score (HSS)} = \frac{2 * (h * cn - fa * m)}{(h + m) * (m + cn) + (h + fa) * (fa + cn)} \quad (1)$$

where h is the hits (both prediction and reference detect an event), cn is the correct negative (prediction and reference do not detect an event), m is the misses (prediction does not detect any event, but reference does) and fa is a false alarm (prediction detects an event, but reference does not). A high POD means that a high proportion of events were correctly predicted, while a high FAR means that a high proportion of predicted events did not occur. HSS combines h, cn, m and fa to assess the prediction skill.

The SWP retrieval module was evaluated using the Pearson correlation coefficient (r), relative bias (Bias) and root mean square error (RMSE) between the predicted and reference datasets. We also computed the SWP fractional standard error percentage (FSE%), which is defined as follows:

$$\text{FSE\%} = \frac{100 * \text{RMSE}}{< \text{SWP} >} \quad (2)$$

where <SWP> is the mean SWP from 2CSP.

Each module has been evaluated independently of the results of the others, e.g., for the supercooled detection module, we considered cases in which snow is detected by the 2CSP, rather than using the outputs of the snow detection module. This assessment strategy ensures that the skill evaluation of each module does not include errors caused by other modules. Obviously, in the operational version of the algorithm (and for the case studies presented in this study), the outputs of the three modules are combined. Therefore, the full SLALOM algorithm, combining the results of the 3 modules, is also evaluated (Section 5.4). The evaluation has been conducted for the whole test dataset and for the following surface types: Open sea, sea ice (sea ice concentration > 90%) and land. Land pixels in the coincident dataset are frozen or snow covered in 88% of the snowfall cases, according to T2m from the GPM-CPR dataset.

5. Results

5.1. Snowfall Detection Module

The snowfall detection module, evaluated on the test dataset, reveals a POD of 0.83, with a FAR of 0.12 and a HSS of 0.84. These statistics are rather similar among surface types: A POD of 0.83, 0.80 and 0.87, a FAR of 0.1, 0.14 and 0.1, and an HSS of 0.85, 0.82 and 0.84 for land, open sea and sea ice, respectively. These results highlight that the performance of the snow detection module does not depend on the surface type.

The sensitivity of the detection module on the 2CSP SWP magnitude is evaluated in Figure 3a. In this Figure, the test dataset is binned as a function of 2CSP SWP and, for each bin, the ratio of the number of missed snow events to the total number of snow events is computed. The percentage of missed snowfall events ranges from 55%, for the lowest SWP given by the 2CSP product (i.e., 3.7×10^{-5} kg·m^{-2}), to less than 10% above 0.02 kg·m^{-2}. We conducted a similar analysis to evaluate the sensitivity of the snowfall detection module to the surface snowfall rate (not shown). It was revealed that the snow detection module misses 58% of snowfall events, when the surface snowfall rate is about 1×10^{-3} mm/h, and less than 10% when the surface snowfall rate exceeds 0.05 mm/h. These findings highlight the high sensitivity of the snow detection module, even for light snowfall events. In order to characterize the cases for which the SLALOM snowfall detection module fails, we computed a decision (regression) tree (Figure 3b). This tool allowed us to identify the environmental variables that explain the percentage of misses by hierarchically partitioning the observations. Thus, in Figure 3b, each leaf (box) contains only a fraction of observations, except for the top box, which contains all the observations. Labels on leaves represent the mean percentage of misses, and the names of variables and inequalities written on branches (black lines) indicate the variable and thresholds chosen to split observations. For each leaf, the tree algorithm selected the most appropriate variables among T2m, TPW, SWP, surface type, and the presence of supercooled droplets for partitioning observations. Thus, in Figure 3b, the average percentage of misses of all observations is 17.1% (top box), and the first variable chosen by the tree to split the observations is SWP. We ensured that each final leaf (bottom box) of the tree contained at least 300 observations in order to obtain reliable results. Results show that the percentage of missed snowfall occurrence reaches 41.4% when SWP is below 3×10^{-3} kg·m^{-2} while it is 10.3% when SWP is above 3×10^{-3} kg·m^{-2}. In addition, the percentage of misses is much higher when T2m exceeds 275 K (44.2% versus 8.8%), as shown in Figure 3b. This can be explained by the fact that when T2m is greater than 273 K, the occurrence of mixed precipitation increases and, when the liquid fraction becomes greater than 15%, the 2CSP product sets the snowfall rate to zero. However, in these cases, GMI channels are still sensitive to melting snow, and this could explain SLALOM detection difficulties in these situations.

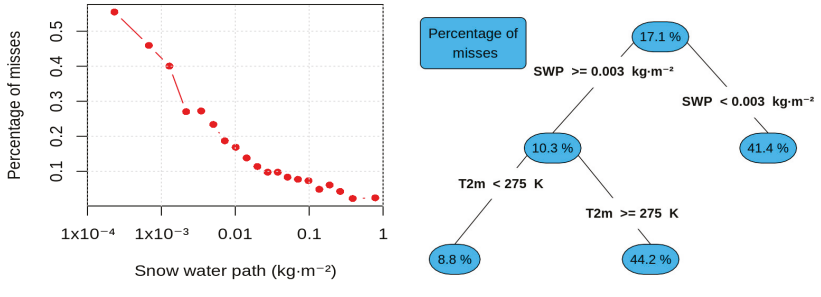


Figure 3. (**a**) Percentage of missed snowfall occurrence as a function of SWP; and (**b**) decision tree, with the percentage of missed snowfall occurrences indicated in boxes, and partitioning variables and thresholds indicated on branches.

5.2. Supercooled Droplets Detection Module

The supercooled detection module shows a POD of 0.97, an FAR of 0.05 and an HSS of 0.89. Once again, the results are rather similar for each surface type: For land (open sea, sea ice), the POD is 0.95 (0.98, 0.97), the FAR is 0.07 (0.04, 0.05) and the HSS is 0.87 (0.9, 0.9). These results illustrate the skill of the random forest approach in distinguishing between snowfall cases with and without supercooled droplets.

The decision tree of the missed and false positive supercooled detection is shown in Figure 4. Each final leaf contains at least 300 observations, and the percentage of missed and false positive supercooled detection is partitioned as a function of T2m, TPW, SWP and surface type. The most significant variable chosen to explain the percentage of missed supercooled droplets is TPW. When TPW is below 2.4 kg·m^{-2}, the module fails to detect supercooled droplets in 11% of cases. These results, unsurprisingly, show that very dry conditions complicate supercooled droplet detection.

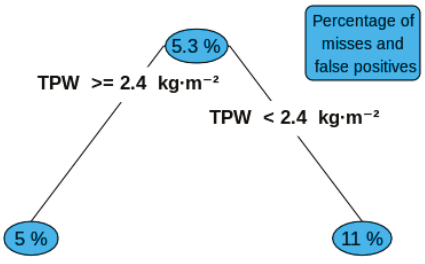


Figure 4. Decision tree, with the percentage of missed and false positive supercooled occurrence indicated in boxes, and partitioning variables and thresholds indicated on branches.

5.3. Snow Retrieval

The skill of the retrieval algorithm, globally and for each surface type, is shown in Table 1. The correlation between the predicted and observed SWP is 0.88, the relative bias is −16% and the RMSE is 0.10 kg·m^{-2}. The results are better over sea ice (r = 0.92, Bias = −15% and RMSE = 0.08 kg·m^{-2}) than over open sea and land surfaces (r = 0.88 and 0.85, Bias = −21% and −13%, and RMSE = 0.12 and 0.1 kg·m^{-2}).

Table 1. Comparison of SWP observations from CPR, with SWP predictions for the complete evaluation dataset (All) over land, open sea and sea ice.

Surface	Correlation	Bias	RMSE
All	0.88	−16%	0.1 kg·m^{-2}
Land	0.85	−13%	0.1 kg·m^{-2}
Open Sea	0.88	−21%	0.12 kg·m^{-2}
Sea ice	0.92	−15%	0.08 kg·m^{-2}

We also investigated the sensitivity of the relative bias and fractional standard error percentage (FSE%) to SWP (Figure 5). The scores have been computed by binning observations as a function of SWP. The relative bias decreases from about 20%, for low SWP (<0.02 kg·m^{-2}), to about −30%, for SWP of 0.05 kg·m^{-2}, and then remains around −20% for SWP higher than 0.05 kg·m^{-2}. Once again, the results are better over sea ice, with a relative bias constantly between −10 and −20%, while it decreases from 40% to −40% for open sea. The FSE% decreases from 200% at 0.01 kg·m^{-2} to about 40% at 1 kg·m^{-2}. It shows a similar pattern for every surface type below 0.05 kg·m^{-2}, and, for SWP over 0.05 kg·m^{-2}, the prediction over open sea shows higher values (around 40% at 1 kg·m^{-2}) compared to sea ice (30% at 0.5 kg·m^{-2}). Land has an intermediate behavior.

A decision tree, with the absolute value of the difference between predicted and observed SWP, normalized by the observed SWP, is shown in Figure 6. Every final leaf contains at least 300 observations, and partitioning variables are T2m, TPW, presence of supercooled droplets and surface type. Note that we retained only cases for which SWP is greater than 0.01 kg·m^{-2}. First, the normalized difference is overall lower when no supercooled droplets are present (i.e., left part of the tree). In addition, if T2m is lower than 273 K, and TPW is higher than 6.7 kg·m^{-2}, the average

normalized difference is minimized (34%). These conditions are very favorable for snow retrievals, since the snow scattering signal is not attenuated by the supercooled droplets. Furthermore, temperatures below freezing point prevent the presence of melting snow and rain that can contaminate the snow radiative signature, while higher TPW reduces any surface contamination. On the right part of the tree, which contains cloud top supercooled droplet occurrences, the worst case is found when TPW is high (94% of normalized difference percentage). It is somewhat surprising that, when no supercooled droplets are present, a high humidity is favorable for accurately retrieving snow, while when supercooled droplets are present, a high humidity is less favorable to the retrieval of snow. This TPW and supercooled water relationship might be explained by the fact that snow events with supercooled droplets are usually associated with lower SWP values. Indeed, the median value of SWP when supercooled droplets occur (about 0.01 $kg \cdot m^{-2}$) is more than ten times lower than without supercooled droplet occurrence (about 0.12 $kg \cdot m^{-2}$), as shown in Figure 6b. Therefore, under these conditions, detecting weak snowfall could be more difficult, because water vapor and cloud droplet emission obscure the weak scattering signal.

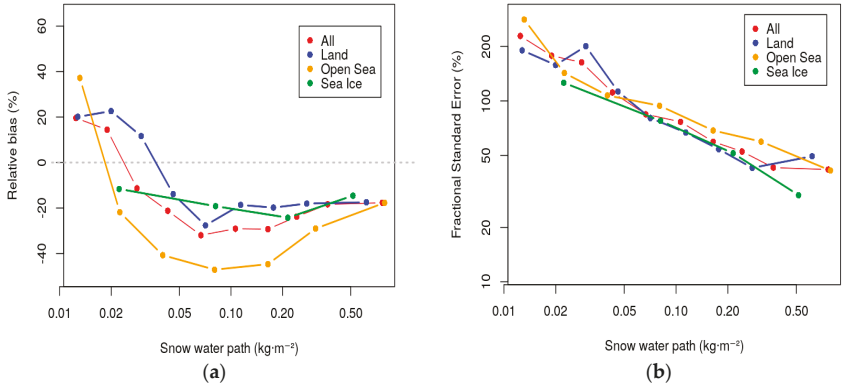


Figure 5. (**a**) Relative bias percentage as a function of SWP and surface type; and (**b**) Fractional Standard Error percentage as a function of SWP and surface type.

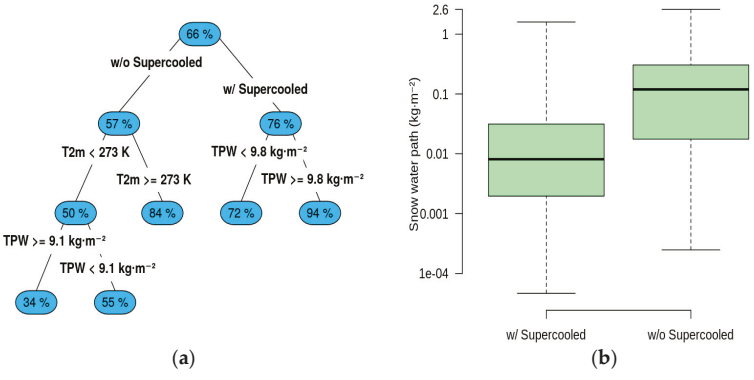


Figure 6. (**a**) Decision tree, with the absolute value of the normalized difference percentage between predicted and observed SWP indicated in boxes, and partitioning variables and thresholds indicated on branches; and (**b**) Boxplot of observed SWP, with supercooled droplets (left) and without supercooled droplets (right).

5.4. Full Algorithm Evaluation and Sensitivity Test

In the previous sections we evaluated each module output independently. In this section, we evaluate the full SLALOM algorithm, combining the results of the 3 modules. Additionally, we included the embedded supercooled droplet cases in this evaluation (while they have not been used for the algorithm training). We also analyzed the sensitivity of the model by evaluating its performance in two experiments: (i) Without supercooled droplet detection and (ii) without using environmental conditions (i.e., only brightness temperatures).

In comparison to observations, global predictions have a correlation of 0.86, a relative bias of -20% and RMSE of 0.04 kg·m^{-2}, as shown in Table 2. Supercooled droplet detection does not have an effect on predictions, meaning that even if supercooled droplets can be detected accurately by SLALOM (Section 5.2), it does not improve the overall SWP retrieval. This could mean that the SWP retrieval module is not able to precisely retrieve SWP when supercooled droplets occur, even if they are appropriately flagged by the supercooled detection module. This interpretation is supported by the fact that the correlation and relative bias (between observations and predictions) are 0.73 and -32% when supercooled droplets are present but reach 0.86 and -16% when no supercooled droplets are present. The fact that supercooled droplets are often associated with a low SWP (Figure 6b) could explain the difficulty of fitting the model to these cases. The importance of considering environmental variables for precisely retrieving SWP is revealed in Table 2. If the model is trained without using environmental conditions, the correlation decreases to 0.67, the relative bias increases to -49% and RMSE increases to 0.13 kg·m^{-2}.

Table 2. Comparison of SWP observations from CPR, with SWP predictions for the full SLALOM algorithm (SLALOM), without taking into account supercooled droplets occurrence (SLALOM w/o Sc) and environmental conditions (SLALOM w/o Env).

Configuration	Correlation	Bias	RMSE
SLALOM	0.86	-20%	0.04 kg·m^{-2}
SLALOM w/o Sc	0.86	-18%	0.04 kg·m^{-2}
SLALOM w/o Env	0.61	-49%	0.13 kg·m^{-2}

6. Applications of SLALOM Algorithm

In the two following sections, we highlight some applications of the SLALOM algorithm. First, we used SLALOM to retrieve the snow water path in the three case studies, described in detail in Panegrossi et al. [43]. These case studies have been chosen by Panegrossi et al. [43], since they represent three distinct and somewhat complex snow situations that can be found at around 60° latitude. They allowed us to evaluate the skill of the algorithm in cases of complex meteorological and environmental conditions. Secondly, we built a $0.1° \times 0.1°$ map of snowfall occurrence from May 2014 to May 2016 in order to illustrate the potential of the SLALOM algorithm for climatological analyses.

6.1. Case Studies

The first case study took place on 30 April 2014 in Eastern Siberia and is associated with an extended frontal system. A maximum SWP of 1.4 kg·m^{-2} is estimated along the CloudSat track (Figure 7c). The frontal structure (around 60°N) is well captured by the snow and supercooled detection modules (Figure 7a). This reveals that the front is composed of a snowy region without supercooled droplets (in red) and, on the southern part of this region, a snowy region with supercooled droplets. The evaluation of the predictions along the CloudSat track shows that the snow detection module performs well in 85% of cases (here 85% is the ratio of the number of hits and correct negatives to the total number of cases) and, in cases where snowfall has been detected successfully, the supercooled detection module was correct in 90% of cases (i.e., the sum of cases with true supercooled droplets detection and true snowfall without supercooled droplets detection). The SWP module predicts values

up to 1.89 kg·m^{-2}, with the most intense values in the southeastern part of the front (Figure 7b). SWP predictions also show very weak patches (<0.1 kg·m^{-2}) on the northern part of the front. SWP observations and predictions match very well in terms of the pattern and magnitude along the CloudSat track, as highlighted in Figure 7c. We also included the ice water path from GPROF (IWP) in Figure 7c. Note that a quantitative comparison with 2CSP or SLALOM is not possible here, since SWP and IWP are different variables (see the conclusion for details). In addition, GPROF does not retrieve IWP over snow covered surfaces (e.g., above 60°N in this case). The highest values of IWP are found where snowfall is detected by 2CSP (with a maximum of 0.25 kg·m^{-2} at around 59.5°N).

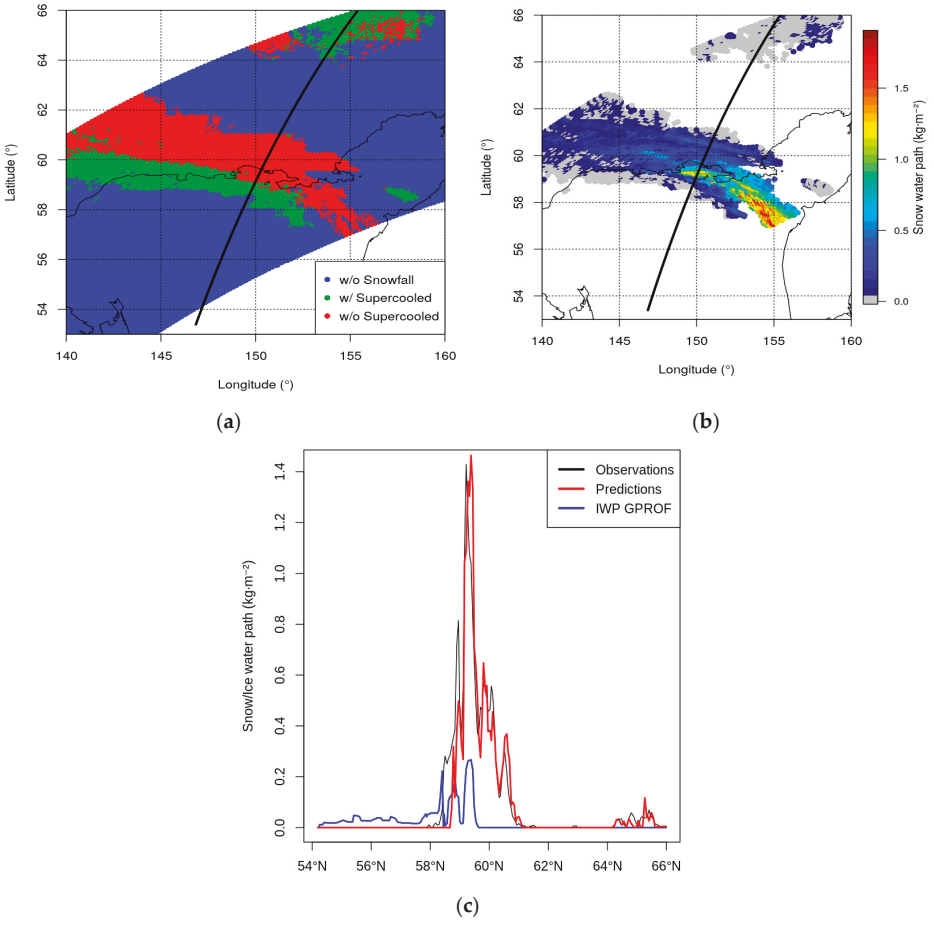


Figure 7. Case study of 30 April 2014: (**a**) Supercooled droplet classification, predicted by SLALOM (blue: Without snow; green: With snowfall and supercooled droplets; red: With snowfall and without supercooled droplets) and the CloudSat track in black; (**b**) retrieved SWP and the CloudSat track in black; and (**c**) SWP cross section along the CloudSat track for observations (black), SLALOM predictions (red), and the GPROF ice water path (IWP) (blue).

The second case study is an orographic precipitation event that occurred on 14 December 2014 in Alaska (Figure 8). Unfortunately, the CloudSat track does not cross the most intense part of this snow event, and very low values of SWP are estimated by 2CSP. Predictions show snowfall occurrences

with and without supercooled droplets along the coast and in the northern region. SWP predictions show some large values (more than 2 kg·m^{-2}) on the eastern part of the region, along the coast (Figure 8b). The evaluation of predictions along the CloudSat track show that, even if the SLALOM algorithm successfully detects snow in the coastal region (around 61°N), predicted SWP does not match observations well and, in addition, several weak SWP peaks (<0.2 kg·m^{-2}) are predicted in the north of the region but are not observed. The bad performance of the algorithm in this case is probably related to extremely dry conditions, with the very low TPW observed inland (lower than 5 kg·m^{-2}), and to the very weak scattering signal, associated with the light snowfall along the CloudSat track, as evidenced by Panegrossi et al. [43]. It is worth noting, however, the ability of SLALOM to provide SWP estimates over the full GMI swath, including the intense snowfall region that is missed by CloudSat. IWP shows a peak of 0.6 kg·m^{-2} at 60°N, while SWP is zero. This could mean that an ice cloud occurred in this region but did not produce snowfall. For latitudes higher than 61°N, IWP is not retrieved because the ground is snow-covered.

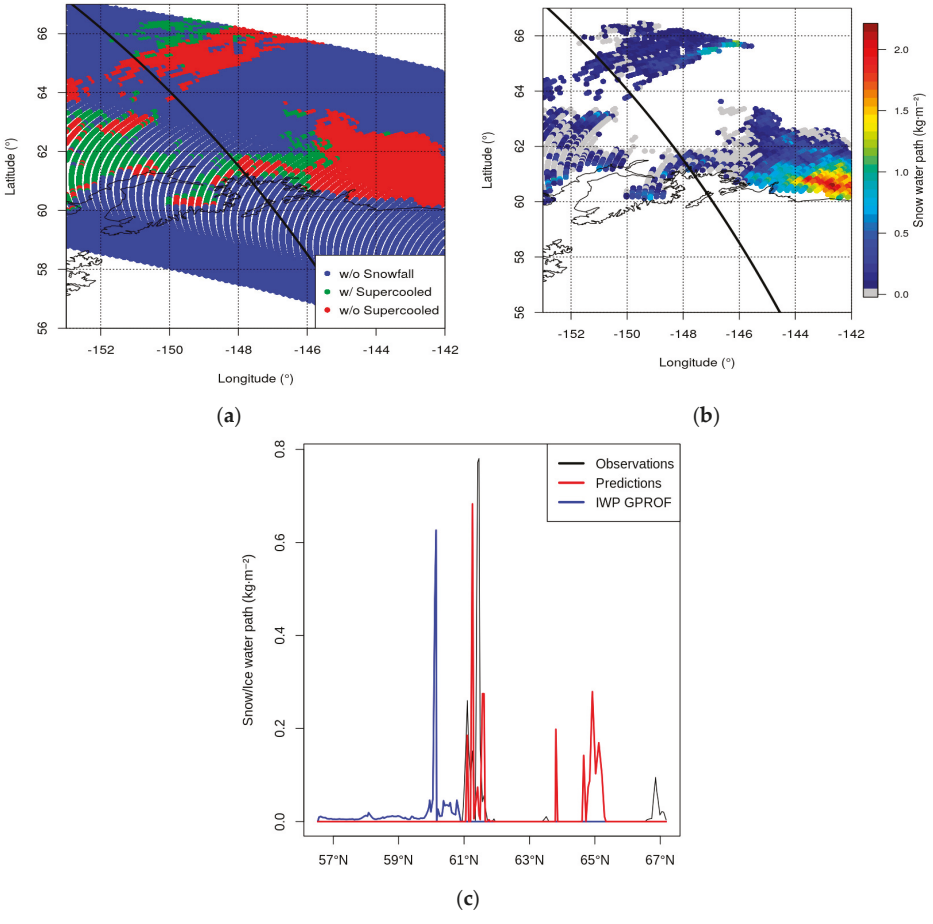


Figure 8. Same as for Figure 7 but for the 14 December 2014 event.

The last investigated case is a synoptic snowfall event that occurred over the Labrador Sea on 27 March 2014. This case is characterized by the transition from open sea to sea ice at around 61°N–58°W.

A large region of snowfall is predicted south of 61°N nearly without the occurrence of supercooled droplets (Figure 9a). The snow and supercooled detection modules perform very well (87% and 94% of good classification). SWP predictions show a large region with moderate SWP values ranging from 0.5 to 1.3 kg·m^{-2} on the eastern part of the region and weaker SWP elsewhere (Figure 9b). The SWP retrieval module also performs well in this case, as it reproduces rather accurately the pattern given by CPR observations. IWP decreases from 0.2 kg·m^{-2} at 57°N to 0 kg·m^{-2} for latitudes above 61°N.

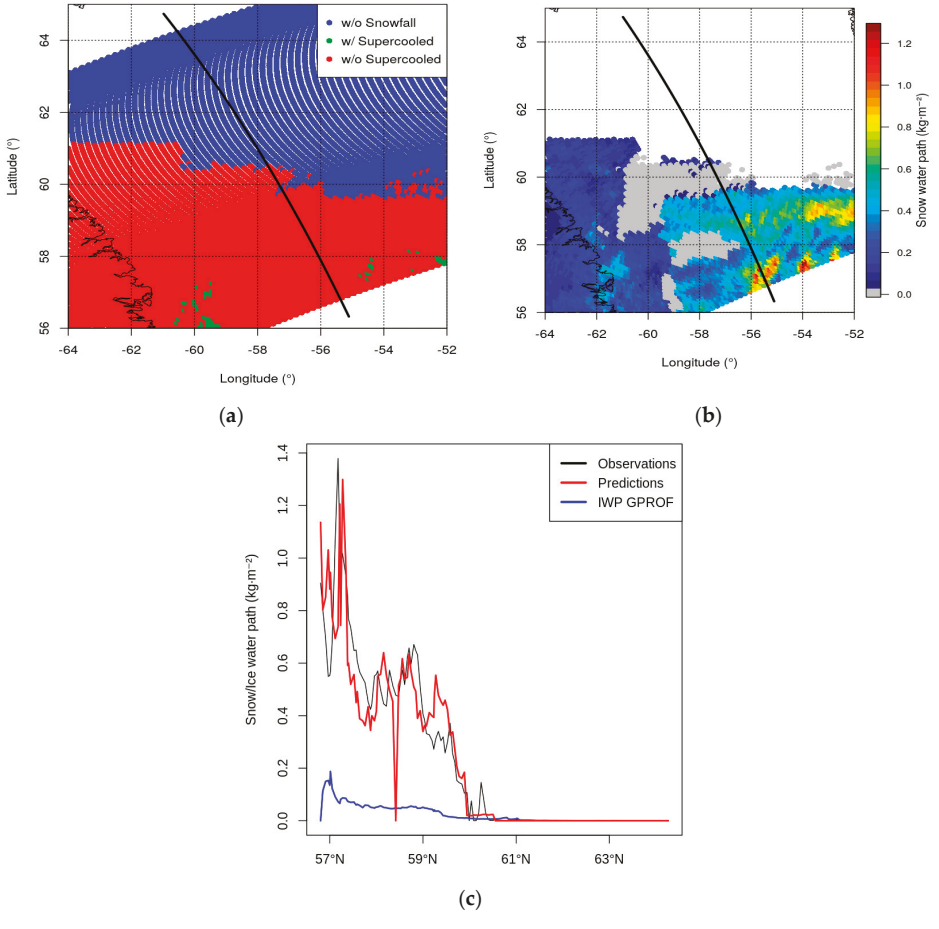


Figure 9. Same as for Figure 7 but for the 27 March 2014 event.

6.2. Climatology of Snowfall Occurrence

The snowfall percentage of occurrence, as given by the snow detection module of SLALOM from May 2014 to May 2016, is shown in Figure 10. For each of GMI orbits between those dates, we applied the snowfall detection module and identified snowfall cases. Then, we binned these detections on a 0.1° × 0.1° grid and normalized it by the total number of GMI observations in each pixel of this grid. For computing optimization reasons, we used here a simplified version of the algorithm to build this climatology, i.e., we did not use vertical profiles of atmospheric variables in the snowfall detection module. The Antarctica coast is the region with the highest occurrence of snowfall, with values reaching 40% in some areas. It is interesting to note the step decrease of snow occurrence as one moves

away from the Antarctica coast. In the Northern Hemisphere, the situation is more complex due to the presence of continental regions. The maximum is found over Greenland, with values over 40% and values of about 30% in the Labrador Sea. Siberia, Canada and the eastern side of continents show an occurrence between 2 and 20%, with local variability. Europe and the western side of continents have a snow occurrence lower than 5% of time. Finally, mountain ranges, such as the Himalayas, the Andes, the Alps and the Rocky Mountains show occurrences that can exceed 20%. These patterns match very well those identified at a lower resolution by Kulie et al. [11] and Behrangi et al. [4], using CloudSat CPR observations, and Adhikari et al. [56], using a GPM Dual-frequency Precipitation Radar.

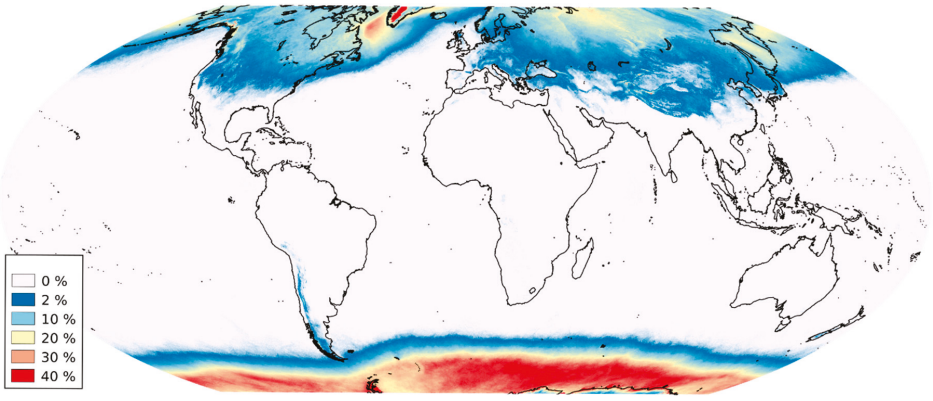


Figure 10. Snowfall percentage of occurrence, as given by the snow detection module of SLALOM, from May 2014 to May 2016 on a 0.1° × 0.1° grid.

7. Discussion and Conclusions

In this paper, we described and evaluated a new algorithm, named SLALOM (Snow retrievaL ALgorithm fOr gMi), that is able to detect snowfall and to retrieve the associated snow water path (SWP). It is tuned for the GPM Microwave Imager (GMI) but can be adapted to other passive microwave radiometers. SLALOM is trained and evaluated using coincident measurements by the CloudSat Cloud Profiling Radar (CPR) and CALIPSO Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). This algorithm is composed of 3 modules: Snowfall detection, supercooled droplet detection and SWP retrieval. It is designed to take into account the impact of environmental conditions (including the occurrence of supercooled droplets) on GMI sensitivity to snowfall and thus to obtain an optimal estimation of SWP. Evaluation of the algorithm reveals that all modules have good skill. Specifically, the snow detection POD is 0.83, with a FAR of 0.12, and supercooled detection shows a POD of 0.97, with a FAR of 0.05. Overall, SWP retrieval shows a correlation of 0.86, a relative bias of −18% and a root mean square error of 0.04 $kg \cdot m^{-2}$. In addition, results show that the SLALOM algorithm is able to detect light snowfall, even if it still misses 55% of the lightest events. We also showed that SLALOM is less efficient in cases of mixed phase precipitation, supercooled droplet occurrence, low humidity and low SWP, and we identified the essential role of environmental variables for a meaningful retrieval of SWP. We applied the SLALOM algorithm to three different meteorological situations (frontal snowfall, orographic snowfall and synoptic snowfall). The algorithm performs very well in the first and the third situations, but it has more difficulty in retrieving snowfall in the orographic snowfall situation, which was characterized by dry conditions and a low SWP, according to 2CSP. Finally, we computed a global climatology of snowfall occurrence from May 2014 to May 2016, with a 0.1° × 0.1° resolution. This shows a very good consistency with lower resolution climatologies using CloudSat CPR [4,11].

Overall, these results highlight the high efficiency of the SLALOM algorithm for every surface type. In particular, the high sensitivity of snow detection by the SLALOM algorithm, even for light

snowfall, is an important asset, as many studies pointed out that most snow events are associated with very light snowfall rates at high latitudes [9,11,13]. In addition, the snowfall and supercooled detection modules can be exploited to document the water phase at the global scale. As illustrated in this paper, the algorithm outputs can be useful for climatic studies related to the water cycle as well as case study analyses, especially in regions where in-situ measurements are not available.

Interestingly, our results illustrate the limited impact of supercooled droplet detection on SWP retrieval accuracy. The reason might be that the supercooled droplet occurrence is often associated with a low SWP, as shown in Section 5.3. Thus, even if the SLALOM algorithm does not perform as well in supercooled droplet cases, the overall error remains small. Therefore, the accurate detection of supercooled mixed phase clouds for SWP retrieval might not be as crucial as it was initially thought to be. In the future, it could also be worthwhile to evaluate the minimum SWP detectable for every GMI channel when supercooled droplets are present, following work of Kneifel et al. [40]. To this end, we could perform idealized radiative transfer simulations of mixed phase clouds and identify the SWP threshold for which the scattering effects exceed cloud top supercooled droplet emission.

Several limitations of the SLALOM algorithm need to be highlighted. First, SLALOM fully relies on the 2C-SNOW-PROFILE (2CSP) product and is therefore subject to the limitations of the 2CSP product. The main limitation is due to the fact that 2CSP does not provide measurements close to the surface. Thus, it misses an important layer of snowfall and underestimates the total SWP. On the other hand, ground clutter is sometimes not appropriately corrected in complicated terrain (see [9,13]), thus leading to artificially high snow content. Another important limitation lies in the fact that 2CSP SWP is subjected to high uncertainties due to the difficulties in quantifying SWP using a single frequency instrument, such as CPR. Indeed, CPR does not provide enough independent information on snow microphysics to be able to connect measured reflectivity to a unique snowfall rate. Thus, the inversion problem is not fully constrained, which necessitates the use of parameterizations in estimating the snowfall rate from measured reflectivity, implying uncertainties in snowfall estimates [57]. Another difficulty comes from melting snow events. When the liquid fraction is higher than 15%, 2CSP does not retrieve snowfall (these are therefore considered no-snowfall cases). However, the high frequency GMI channels are affected by melting snowfall, even if the liquid fraction is higher than 15%, which can create discrepancies between both SLALOM and 2CSP. In addition, coincidences between CPR and GMI are unevenly distributed (Figure 1), which can also affect the overall predictions.

Secondly, the SLALOM algorithm design itself has some limitations. For instance, the snow retrieval module assumes a segmented multi-linear relationship between high frequency brightness temperatures and SWP. It is probable that a more realistic (but complex) relationship could be found and could reduce SWP retrieval uncertainties. In addition, we only trained the algorithm using supercooled droplets on top of the cloud and discounted supercooled droplet layers embedded within the cloud. However, embedded supercooled layers frequently occur in the coincident database (about 30% of time). Depending on the position in the cloud, supercooled droplet layers can mask more or less of the snow scattering signal. Considering the embedded cases could improve the algorithm predictions.

Previous studies on the GMI snowfall detection capability assessment [18,26,33,43] demonstrate that GMI has great potential for snowfall observation. In this work, we have shown that, through the exploitation of all 13 GMI channels and the optimal use of ancillary variables describing the atmospheric conditions (and no ancillary information on the background surface conditions), SLALOM is able to predict snowfall occurrence and SWP, with a very good agreement with the Cloudsat 2CSP product, and has the advantage of ensuring a much larger spatial coverage, corresponding to the GMI swath. In spite of some of the limitations already mentioned, SLALOM can be applied effectively in case studies as well as in climatology analysis. SLALOM will be further developed to provide the surface snowfall rate and incorporated in the global precipitation retrieval algorithm for GMI, developed recently within the EUMETSAT Satellite Application Facility on support to Operational Hydrology and Water Management (H SAF) program [58]. In the future, studies similar to that of

Panegrossi et al. [43], based on the use of experimental datasets built from coincident observations by Cloudsat and other passive microwave radiometers in the GPM constellation (conical and cross-track scanning), will be carried out, and approaches similar to SLALOM will be developed to extend snowfall detection and retrieval to higher latitudes, and they will have a global coverage of snowfall by the passive microwave radiometer constellation. In addition, a comparison of the SLALOM product with other operational products, such as GPROF, will be conducted. There are some important factors to be considered in order to conduct such comparison. One is the difficulty in finding high-quality and independent snowfall observations and estimates to be used as the "truth". Then, there are product differences, often difficult to reconcile, that can derive from different sources: the method used to determine the phase of the precipitation on the surface; the differences in orbits and swath widths; instrumentation differences (channel assortment and resolution); and differences in the algorithm formulation and assumptions, which are oftenf not well understood [59]. Review papers, similar to that of Levizzani et al. [3], summarizing the main snowfall retrieval techniques currently available, are thus strongly needed.

Author Contributions: J.F.R. designed and implemented the SLALOM algorithm. J.F.R. and G.P. wrote this paper. All co-authors have contributed to the group discussions on the development of the algorithm and on the results, and to the final draft.

Funding: This work is supported by the EUMETSAT H SAF [CDOP-2 and CDOP-3]. Anna Cinzia Marra is supported by the Italian Research Project of National Interest 2015 (PRIN 2015) [4WX5NA]. Mark Kulie and Lisa Milani, NASA Grant [NNX16AE21G] is also gratefully acknowledged.

Acknowledgments: The datasets used in this study have been collected from the NASA PPS website ftp://arthurhou.pps.eosdis.nasa.gov (GPM products, 2B-CSATGPM product), from the CloudSat website ftp.cloudsat.cira.colostate.edu (CPR products), from http://www.icare.univ-lille1.fr/projects/dardar (DARDAR product), and from https://seaice.uni-bremen.de/sea-ice-concentration/ (AMSR2 sea ice dataset). The PMM Research Program and EUMETSAT are warmly acknowledged for supporting the H SAF and GPM collaboration through the approval of the no-cost H SAF-GPM proposal, "H SAF and GPM: Precipitation algorithm development and validation activity". The authors want to sincerely express their gratitude to Joe Turk (NASA JPL), for developing and sharing the 2B-CSATGPM dataset used in this study, and to Norm Wood, for the valuable discussions and suggestions on the use of the CloudSat 2C-SNOW-profile product. We would also like to thank the 4 referees of this paper for their valuable comments.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, and interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Tools: This project has been conducted using the R software (version 3.4 [60]), with the following packages: rhdf5 [61], ncdf4 [62], rpart [63], data.table [64], randomForest [65], caret [66], party [67], partykit [68], lme4 [69], MASS [70], hydroGOF [71], tree [72], nabor [73], snow [74], snowfall [75], segmented [76,77], rgl [78].

References

1. Knowles, N.; Dettinger, M.D.; Cayan, D.R. Trends in snowfall versus rainfall in the Western United States. *J. Clim.* **2006**, *19*, 4545–4559. [CrossRef]
2. Feng, S.; Hu, Q. Changes in winter snowfall/precipitation ratio in the contiguous United States. *J. Geophys. Res. Atmos.* **2007**, *112*, D15109. [CrossRef]
3. Levizzani, V.; Laviola, S.; Cattani, E. Detection and measurement of snowfall from space. *Remote Sens.* **2011**, *3*, 145–166. [CrossRef]
4. Behrangi, A.; Christensen, M.; Richardson, M.; Lebsock, M.; Stephens, G.; Huffman, G.J.; Bolvin, D.; Adler, R.F.; Gardner, A.; Lambrigtsen, B.; et al. Status of high-latitude precipitation estimates from observations and reanalyses. *J. Geophys. Res.-Atmos.* **2016**, *121*, 4468–4486. [CrossRef] [PubMed]
5. Skofronick-Jackson, G.M.; Kim, M.J.; Weinman, J.A.; Chang, D.E. A physical model to determine snowfall over land by microwave radiometry. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1047–1058. [CrossRef]
6. Skofronick-Jackson, G.M.; Johnson, B.T.; Munchak, S.J. Detection thresholds of falling snow from satellite-borne active and passive sensors. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4177–4189. [CrossRef]

7. Stephens, G.L.; Vane, D.G.; Boain, R.J.; Mace, G.G.; Sassen, K.; Wang, Z.E.; Illingworth, A.J.; O'Connor, E.J.; Rossow, W.B.; Durden, S.L.; et al. The cloudsat mission and the a-train—A new dimension of space-based observations of clouds and precipitation. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1771–1790. [CrossRef]
8. 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]
9. Kulie, M.S.; Bennartz, R. Utilizing spaceborne radars to retrieve dry snowfall. *J. Appl. Meteorol. Climatol.* **2009**, *48*, 2564–2580. [CrossRef]
10. Hiley, M.J.; Kulie, M.S.; Bennartz, R. Uncertainty analysis for CloudSat snowfall retrievals. *J. Appl. Meteorol. Climatol.* **2011**, *50*, 399–418. [CrossRef]
11. Kulie, M.S.; Milani, L.; Wood, N.B.; Tushaus, S.A.; Bennartz, R.; L'Ecuyer, T.S. A shallow cumuliform snowfall census using spaceborne radar. *J. Hydrometeorol.* **2016**, *17*, 1261–1279. [CrossRef]
12. Chen, S.; Hong, Y.; Kulie, M.; Behrangi, A.; Stepanian, P.M.; Cao, Q.; You, Y.; Zhang, J.; Hu, J.; Zhang, X. Comparison of snowfall estimates from the NASA CloudSat cloud profiling radar and NOAA/NSSL multi-radar multi-sensor system. *J. Hydrol.* **2016**, *541*, 862–872. [CrossRef]
13. Milani, L.; Kulie, M.S.; Casella, D.; Dietrich, S.; L'Ecuyer, T.S.; Panegrossi, G.; Porcù, F.; Sanò, P.; Wood, N.B. CloudSat snowfall estimates over Antarctica and the Southern Ocean: An assessment of independent retrieval methodologies and multi-year snowfall analysis. *Atmos. Res.* **2018**, *213*, 121–135. [CrossRef]
14. Casella, D.; Panegrossi, G.; Sano, P.; Marra, A.C.; Dietrich, S.; Johnson, B.T.; Kulie, M.S. Evaluation of the GPM-DPR snowfall detection capability: Comparison with CloudSat-CPR. *Atmos. Res.* **2017**, *197*, 64–75. [CrossRef]
15. Bennartz, R.; Bauer, P. Sensitivity of microwave radiances at 85–183 GHz to precipitating ice particles. *Radio Sci.* **2003**, *38*. [CrossRef]
16. Liu, G.; Seo, E.-K. Detecting snowfall over land by satellite high-frequency microwave observations: The lack of scattering signature and a statistical approach. *J. Geophys. Res.-Atmos.* **2013**, *118*, 1376–1387. [CrossRef]
17. Skofronick-Jackson, G.; Johnson, B.T. Surface and atmospheric contributions to passive microwave brightness temperatures for falling snow events. *J. Geophys. Res.-Atmos.* **2011**, *116*, D02213. [CrossRef]
18. Gong, J.; Wu, D.L. Microphysical properties of frozen particles inferred from Global Precipitation Measurement (GPM) Microwave Imager (GMI) polarimetric measurements. *Atmos. Chem. Phys.* **2017**, *17*, 2741–2757. [CrossRef]
19. Liu, G.S.; Curry, J.A. Precipitation characteristics in Greenland-Iceland-Norwegian Seas determined by using satellite microwave data. *J. Geophys. Res.-Atmos.* **1997**, *102*, 13987–13997. [CrossRef]
20. Kongoli, C.; Pellegrino, P.; Ferraro, R.R.; Grody, N.C.; Meng, H. A new snowfall detection algorithm over land using measurements from the Advanced Microwave Sounding Unit (AMSU). *Geophys. Res. Lett.* **2003**, *30*, 1756. [CrossRef]
21. Surussavadee, C.; Staelin, D.H. Satellite retrievals of arctic and equatorial rain and snowfall rates using millimeter wavelengths. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3697–3707. [CrossRef]
22. Noh, Y.-J.; Liu, G.; Jones, A.S.; Haar, T.H.V. Toward snowfall retrieval over land by combining satellite and in situ measurements. *J. Geophys. Res.-Atmos.* **2009**, *114*, D24205. [CrossRef]
23. Kongoli, C.; Meng, H.; Dong, J.; Ferraro, R. A snowfall detection algorithm over land utilizing high-frequency passive microwave measurements-Application to ATMS. *J. Geophys. Res.-Atmos.* **2015**, *120*, 1918–1932. [CrossRef]
24. Kongoli, C.; Meng, H.; Dong, J.; Ferraro, R. A hybrid snowfall detection method from satellite passive microwave measurements and global forecast weather models. *Q. J. R. Meteorol. Soc.* **2018**. [CrossRef]
25. You, Y.; Wang, N.-Y.; Ferraro, R. A prototype precipitation retrieval algorithm over land using passive microwave observations stratified by surface condition and precipitation vertical structure. *J. Geophys. Res.-Atmos.* **2015**, *120*, 5295–5315. [CrossRef]
26. You, Y.; Wang, N.-Y.; Ferraro, R.; Rudlosky, S. Quantifying the snowfall detection performance of the GPM microwave imager channels over land. *J. Hydrometeorol.* **2017**, *18*, 729–751. [CrossRef]
27. Kummerow, C.D.; Randel, D.L.; Kulie, M.; Wang, N.-Y.; 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]
28. Sims, E.M.; Liu, G. A parameterization of the probability of snow–rain transition. *J. Hydrometeorol.* **2015**, *16*, 1466–1477. [CrossRef]

29. Skofronick-Jackson, G.; Munchak, S.J.; Ringerud, S.; Petersen, W.; Lott, B. Falling snow estimates from the global precipitation measurement (gpm) mission. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; IEEE: New York, NY, USA, 2017; pp. 2724–2727.
30. Prigent, C.; Aires, F.; Rossow, W.B. Land surface microwave emissivities over the globe for a decade. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 1573–1584. [CrossRef]
31. Foster, J.L.; Skofronick-Jackson, G.; Meng, H.; Wang, J.R.; Riggs, G.; Kocin, P.J.; Johnson, B.T.; Cohen, J.; Hall, D.K.; Nghiem, S.V. Passive microwave remote sensing of the historic February 2010 snowstorms in the Middle Atlantic region of the USA. *Hydrol. Process.* **2012**, *26*, 3459–3471. [CrossRef]
32. 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.* **2017**. [CrossRef]
33. Ebtehaj, A.M.; Kummerow, C.D. Microwave retrievals of terrestrial precipitation over snow-covered surfaces: A lesson from the GPM satellite. *Geophys. Res. Lett.* **2017**, *44*, 6154–6162. [CrossRef]
34. Petty, G. Physical retrievals of over-ocean rain rate from multichannel microwave imagery. 1: Theoretical characteristics of normalized polarization and scattering indexes. *Meteorol. Atmos. Phys.* **1994**, *54*, 79–99. [CrossRef]
35. 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]
36. Kulie, M.S.; Bennartz, R.; Greenwald, T.J.; Chen, Y.; Weng, F. Uncertainties in microwave properties of frozen precipitation implications for remote sensing and data assimilation. *J. Atmos. Sci.* **2010**, *67*, 3471–3487. [CrossRef]
37. Petty, G.W.; Huang, W. Microwave backscatter and extinction by soft ice spheres and complex snow aggregates. *J. Atmos. Sci.* **2010**, *67*, 769–787. [CrossRef]
38. Kuo, K.-S.; Olson, W.S.; Johnson, B.T.; Grecu, M.; Tian, L.; Clune, T.L.; van Aartsen, B.H.; Heymsfield, A.J.; Liao, L.; Meneghini, R. The microwave radiative properties of falling snow derived from nonspherical ice particle models. Part I: An extensive database of simulated pristine crystals and aggregate particles, and their scattering properties. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 691–708. [CrossRef]
39. Olson, W.S.; Tian, L.; Grecu, M.; Kuo, K.-S.; Johnson, B.T.; Heymsfield, A.J.; Bansemer, A.; Heymsfield, G.M.; Wang, J.R.; Meneghini, R. The microwave radiative properties of falling snow derived from nonspherical ice particle models. Part II: Initial testing using radar, radiometer and in situ observations. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 709–722. [CrossRef]
40. Kneifel, S.; Loehnert, U.; Battaglia, A.; Crewell, S.; Siebler, D. Snow scattering signals in ground-based passive microwave radiometer measurements. *J. Geophys. Res.-Atmos.* **2010**, *115*, D16214. [CrossRef]
41. Wang, Y.; Liu, G.; Seo, E.-K.; Fu, Y. Liquid water in snowing clouds: Implications for satellite remote sensing of snowfall. *Atmos. Res.* **2013**, *131*, 60–72. [CrossRef]
42. Johnson, B.T.; Olson, W.S.; Skofronick-Jackson, G. The microwave properties of simulated melting precipitation particles: Sensitivity to initial melting. *Atmos. Meas. Tech.* **2016**, *9*, 9–21. [CrossRef]
43. Panegrossi, G.; Rysman, J.-F.; Casella, D.; Marra, A.C.; Sano, P.; Kulie, M.S. CloudSat-based assessment of GPM microwave imager snowfall observation capabilities. *Remote Sens.* **2017**, *9*, 1263. [CrossRef]
44. Panegrossi, G.; Rysman, J.-F.; Casella, D.; Sano, P.; Marra, A.C.; Dietrich, S.; Kulie, M.S. Exploitation of GPM/CloudSat coincidence dataset for global snowfall retrieval. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 23–27 July 2018; IEEE: New York, NY, USA, 2018.
45. Draper, D.W.; Newell, D.A.; Wentz, F.J.; Krimchansky, S.; Skofronick-Jackson, G.M. The global precipitation measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2015**, *8*, 3452–3462. [CrossRef]
46. Skofronick-Jackson, G.; Berg, W.; Kidd, C.; Kirschbaum, D.B.; Petersen, W.A.; Huffman, G.J.; Takayabu, Y.N. Global precipitation measurement (GPM): Unified precipitation estimation from space. In *Remote Sensing of Clouds and Precipitation*; Andronache, C., Ed.; Springer International Publishing: Cham, Switzerland, 2018; pp. 175–193. ISBN 978-3-319-72583-3.
47. Turk, J. CloudSat-GPM coincidence dataset (version 1C). In *NASA Technical Report*; California Institute of Technology: Pasadena, CA, USA, 2016.

48. Xie, X.; Loehnert, U.; Kneifel, S.; Crewell, S. Snow particle orientation observed by ground-based microwave radiometry. *J. Geophys. Res.-Atmos.* **2012**, *117*, D02206. [CrossRef]
49. Battaglia, A.; Delanoë, J. Synergies and complementarities of CloudSat-CALIPSO snow observations. *J. Geophys. Res.-Atmos.* **2013**, *118*, 721–731. [CrossRef]
50. 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]
51. Spreen, G.; Kaleschke, L.; Heygster, G. Sea ice remote sensing using AMSR-E 89-GHz channels. *J. Geophys. Res.-Oceans* **2008**, *113*, C02S03. [CrossRef]
52. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
53. Liu, J.; Wu, S.Y.; Zidek, J.V. On segmented multivariate regression. *Stat. Sin.* **1997**, *7*, 497–525.
54. Breiman, L. *Classification and Regression Trees*; Routledge: Abingdon, UK, 2017.
55. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*; Academic Press: Cambridge, MA, USA, 2011; ISBN 978-0-12-385022-5.
56. Adhikari, A.; Liu, C.; Kulie, M.S. Global distribution of snow precipitation features and their properties from 3 years of GPM observations. *J. Clim.* **2018**, *31*, 3731–3754. [CrossRef]
57. Wood, N.B.; L'Ecuyer, T.S.; Heymsfield, A.J.; Stephens, G.L.; Hudak, D.R.; Rodriguez, P. Estimating snow microphysical properties using collocated multisensor observations. *J. Geophys. Res. Atmos.* **2014**, *119*, 8941–8961. [CrossRef]
58. Sano, P.; Panegrossi, G.; Casella, D.; Marra, A.C.; D'Adderio, L.P.; Rysman, J.-F.; Dietrich, S. The passive microwave neural network precipitation retrieval (PNPR) algorithm for the conical scanning GMI radiometer. *Remote Sens.* **2018**, *10*, 1122. [CrossRef]
59. Skofronick-Jackson, G.; Kulie, M.S.; Milani, L.; Munchak, S.J.; Wood, N.B.; Levizzani, V. Satellite estimation of falling snow: A global precipitation measurement (GPM) core observatory perspective. *Rev. J. Appl. Meteorol. Climatol.* (under review).
60. RCoreTeam. *R: A Language and Environment for Statistical Computing*; RCoreTeam: Vienna, Austria, 2013.
61. Fischer, B. rhdf5-HDF5 interface for R. In *R# Package Version*; RCoreTeam: Vienna, Austria, 2015; Volume 2.
62. Pierce, D. ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. R Package 2012. Available online: http://CRAN.R-project.org/package=ncdf4 (accessed on 31 January 2018).
63. Therneau, T.; Atkinson, B.; Ripley, B. RPART: Recursive Partitioning and Regression Trees. In *R Package Version 4.1–10*; RCoreTeam: Vienna, Austria, 2015.
64. Dowle, M.; Short, T.; Lianoglou, S.; Saporta, R.; Srinivasan, A.; Antonyan, E. Data. Table: Extension of Data. Frame. 2014. Available online: https://cran.r-project.org/web/packages/data.table/index.html (accessed on 31 January 2018).
65. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
66. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A. Caret: Classification and regression training. 2016. In *R Package Version*; RCoreTeam: Vienna, Austria, 2017; Volume 4.
67. Hothorn, T.; Hornik, K.; Zeileis, A. Unbiased recursive partitioning: A conditional inference framework. *J. Comput. Graph. Stat.* **2006**, *15*, 651–674. [CrossRef]
68. Zeileis, A.; Hothorn, T.; Hornik, K. Model-based recursive partitioning. *J. Comput. Graph. Stat.* **2008**, *17*, 492–514. [CrossRef]
69. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *arXiv*. 2014. arXiv Preprint:1406.5823. Available online: https://arxiv.org/abs/1406.5823 (accessed on 31 January 2018).
70. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S-PLUS*; Springer Science & Business Media: New York, NY, USA, 2013; ISBN 978-1-4757-3121-7.
71. Zambrano-Bigiarini, M. hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. In *R Package Version 0.3-8*; RCoreTeam: Vienna, Austria, 2014.
72. Ripley, B. Tree: Classification and regression trees. In *R Package Version*; RCoreTeam: Vienna, Austria, 2005. Available online: http://CRAN.R-project.org/package=tree (accessed on 31 January 2018).
73. Elseberg, J.; Magnenat, S.; Siegwart, R.; Nüchter, A. Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration. *J. Softw. Eng. Robot.* **2012**, *3*, 2–12.

74. Tierney, L.; Rossini, A.J.; Li, N.; Sevcikova, H. *Snow: Simple network of workstations. R Package Version 0.3-3*; RCoreTeam: Vienna, Austria, 2008. Available online: http://CRAN.R-project.org/package=snow (accessed on 31 January 2018).
75. Knaus, J. Snowfall: Easier cluster computing (based on snow). In *R Package Version*; RCoreTeam: Vienna, Austria, 2010; Volume 1.
76. Muggeo, V.M. Estimating regression models with unknown break-points. *Stat. Med.* **2003**, *22*, 3055–3071. [CrossRef] [PubMed]
77. Muggeo, V.M. Segmented: An R package to fit regression models with broken-line relationships. *R news* **2008**, *8*, 20–25.
78. Adler, D.; Murdoch, D.; Nenadic, O.; Urbanek, S.; Chen, M.; Gebhardt, A.; Bolker, B.; Csardi, G.; Strzelecki, A.; Senger, A. Rgl: 3D visualization using OpenGL. In *R Package Version 0.95*; RCoreTeam: Vienna, Austria, 2016; Volume 1441.

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Remote Sensing Editorial Office
E-mail: remotesensing@mdpi.com
www.mdpi.com/journal/remotesensing

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-03921-286-6