

Article

An Appraisal of Dynamic Bayesian Model Averaging-based Merged Multi-Satellite Precipitation Datasets Over Complex Topography and the Diverse Climate of Pakistan

Khalil Ur Rahman ¹, Songhao Shang ^{1,*}, Muhammad Shahid ^{1,2} and Yeqiang Wen ¹

- State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China; hlrhm17@mails.tsinghua.edu.cn (K.U.R.);
- m.shahid@nice.nust.edu.pk (M.S.); wenyq15@mails.tsinghua.edu.cn (Y.W.)
 NICE, SCEE, National University of Sciences & Technology Islamabad, Islamabad 44000, Pakistan
- * Correspondence: shangsh@tsinghua.edu.cn; Tel.: +86-10-6279-6674

Received: 28 October 2019; Accepted: 16 December 2019; Published: 18 December 2019



Abstract: Merging satellite precipitation products tends to reduce the errors associated with individual satellite precipitation products and has higher potential for hydrological applications. The current study evaluates the performance of merged multi-satellite precipitation dataset (daily temporal and 0.25° spatial resolution) developed using the Dynamic Bayesian Model Averaging algorithm across four different climate regions, i.e., glacial, humid, arid and hyper-arid regions, of Pakistan during 2000–2015. Four extensively evaluated SPPs over Pakistan, i.e., Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B42V7, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR), Climate Prediction Center MORPHing technique (CMORPH), and Era-Interim, are used to develop the merged multi-satellite precipitation dataset. Six statistical indices, including Mean Bias Error, Mean Absolute Error, Root Mean Square Error, Correlation Coefficient, Kling-Gupta efficiency, and Theil's U coefficient, are used to evaluate the performance of merged multi-satellite precipitation dataset over 102 ground precipitation gauges both spatially and temporally. Moreover, the ensemble spread score and standard deviation are also used to depict the spread and variation of precipitation of merged multi-satellite precipitation dataset. Skill scores for all statistical indices are also included in the analyses, which shows improvement of merged multi-satellite precipitation dataset against Simple Model Averaging. The results revealed that DBMA-MSPD assigned higher weights to TMPA (0.32) and PERSIANN-CDR (0.27). TMPA presented higher skills in glacial and humid regions with average weights of 0.32 and 0.37 as compared to PERSIANN-CDR of 0.27 and 0.25, respectively. TMPA and Era-Interim depicted higher skills during pre-monsoon and monsoon seasons, with average weights of 0.31 and 0.52 (TMPA) and 0.25 and 0.21 (Era-Interim), respectively. Merged multi-satellite precipitation dataset overestimated precipitation in glacial/humid regions and showed poor performance, with the poorest values of mean absolute error (2.69 mm/day), root mean square error (11.96 mm/day), correlation coefficient (0.41), Kling-Gupta efficiency score (0.33) and Theil's U (0.70) at some stations in glacial/humid regions. Higher performance is observed in hyper-arid region, with the best values of 0.71 mm/day, 1.72 mm/day, 0.84, 0.93, and 0.37 for mean absolute error, root mean square error, correlation coefficient, Kling-Gupta Efficiency score, and Theil's U, respectively. Merged multi-Satellite Precipitation Dataset demonstrated significant improvements as compared to TMPA across all climate regions with average improvements of 45.26% (mean bias error), 30.99% (mean absolute error), 30.1% (root mean square error), 11.34% (correlation coefficient), 9.53% (Kling-Gupta efficiency score) and 8.86% (Theil's U). The ensemble spread and variation of DBMA-MSPD calculated using ensemble spread score and standard deviation demonstrates high spread (11.38 mm/day) and variation (12.58 mm/day) during monsoon season in the humid and glacial regions, respectively.



Moreover, the improvements of DBMA-MSPD quantified against fixed weight SMA-MSPD reveals supremacy of DBMA-MSPD, higher improvements (40–50%) in glacial and humid regions.

Keywords: satellite precipitation; merged multi-satellite precipitation dataset; Dynamic Bayesian Model Averaging; regional and seasonal evaluation; complex topography; diverse climate; Pakistan

1. Introduction

Precipitation is one of the integral components in climate studies and global water and energy cycles [1]. Precipitation is used as a forcing data in hydrological modeling such as weather monitoring, streamflow forecasting, flood simulations and warnings, and hydrological trend analysis [2]. Therefore, the accuracy of hydrological simulations and applications is contingent on the accuracy of precipitation. Conventional ground-based precipitation measurements, including rain (or snow) gauges networks and weather radar systems, are either sparsely distributed both spatially and temporally or do not exist at all due to climate variability, complex topography, human geography, and other limited conditions [3,4]. However, ground precipitation gauges (GPGs) are considered the most accurate method for measuring the precipitation because they provide direct precipitation observations [5].

Recent advancements in remote sensing technologies provide an alternative solution to surmount the spatial and temporal limitations of conventional ground-based precipitation measurement networks. Satellite-based precipitation products (SPPs) have been utilized during the past thirty years for regional to global scale precipitation estimation [6,7]. At present, SPPs have been extensively applied in hydrological simulations [8–10], regional and global precipitation pattern recognition, and drought monitoring [11–13]. The applications of SPPs are complicated, and the proper selection of one SPP over another is debatable. Therefore, the evaluation of each SPP against the in-situ observations before the application is extremely important and considered a standard process [14,15].

Various studies have evaluated different SPPs against the GPGs both on regional and global scales at daily [16–22], monthly [17,19,23,24], and annual [17,20] temporal resolutions. It is concluded from the previous studies that all the SPPs have inevitable errors based on different statistical indices, including mean error, correlation with GPGs, and other categorical indices (probability of detection, critical success index, and false alarm ratio). The errors might be associated with precipitation retrieval algorithms or other factors such as the topographic, climatic and seasonal dependency of SPPs. These errors impact the quality of precipitation estimates, which propagate further into hydrological modeling and other applications [16].

Several efforts, such as developing new products with high spatial and temporal resolution [23,25], reducing sampling issues, improvements in precipitation estimation algorithms, merging different SPPs using relative weights and weight estimation using the dynamic methods, have been made to address the uncertainties of existing SPPs [7,16,26]. Developing and application of a merged multi-satellite precipitation dataset (MSPD) based on different statistical models or even the same statistical model can significantly improve the accuracy of meteorological and hydrological models [27]. The performance of MSPD is generally better than all or most of the individual merging members. Several methods are proposed in literature to develop MSPDs, which include Statistical Objective Analysis [28], Conditional Merging [29,30], Simple Scaling Method [30], data assimilation [31], Bayesian Model Averaging [26,27,32], probability density function [33], variation approach [34], and neural network analysis [35]. Furthermore, Li and Shao [36] merged gauged precipitation and TRMM datasets over Australia using a non-parametric kernel merging method, and concluded high performance of the developed MSPD. Readers are referred to [36–39] for more details about techniques to develop MSPDs.

Many studies reported significantly better performance of MSPD developed for estimation of climate and hydrological variables using a simple model averaging (SMA) approach [24,27,40]. Shen et al. [41] evaluated the One-Outlier Removed (OOR) method and reported better performance of

OOR than SMA for all seasons except the winter in Tibetan Plateau. Bayesian Model Averaging (BMA) is a more progressive merging technique, which calculates the optimum weights for individual merging members based on the ground observations. In this method, consensus predictions are derived from multiple competing predictors, and MSPD is generated to exploit the strength of every prediction. BMA method assigns weights to each member based on its predictive performance [27,42,43]. It provides a more convincing portrayal of the predictive uncertainty that considers both within-model and between-model variances and also outperformed other merging techniques by producing reliable and precise results [40,42,43]. This method has also been successfully applied in hydrological modeling and streamflow simulations [44–46]. A MSPD using Dynamic BMA (DBMA) was developed recently to optimally merge four SPPs and evaluated over Tibetan Plateau both topographically and seasonally [26,47]. It accounts for both spatial and temporal variability in the performances of the merging members. The DBMA-MSPD was found superior to both SMA and OOR merging methods [26].

Very limited literature is available on developing and evaluating the MSPDs across Pakistan [16,22,48]. The results demonstrated significant improvements of MSPDs overall merging members. However, most of the literature used the fixed weight techniques [16,22] or the dynamic weight algorithms to merge different SPPs in each climate region [48], and the spatial variation of these weights are not fully considered. The present MSPD considers both spatial and temporal variation in weights for different SPPs over complex topography and diverse climate of Pakistan. The key objectives of the current study are to develop a MSPD using the Dynamic Bayesian Model Averaging (DBMA-MSPD) algorithm for whole Pakistan, to evaluate and compare the performance of DBMA-MSPD against the individual members at seasonal and topographical scales. This experiment is performed using four comprehensively evaluated satellite products in Pakistan, including Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), Climate Prediction Center MORPHing technique (CMORPH), ERA-Interim, and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN-CDR), at daily scale for sixteen years (2000–2015). This study is organized as follows: Section 2 introduces the study area, data and methods, Section 3 represents the comprehensive evaluation and discussion, and Section 4 is the conclusion of the current research.

2. Materials and Methods

2.1. Study Area

Pakistan lies between 23.5° – 37.5° N latitude and 62° – 75° E longitude having an area of 803,940 km² (Figure 1a). Pakistan shares its boundaries with China at its north, India at east, Iran, and Afghanistan at west and the Arabian Sea at south. The study area is topographically very complex, with a maximum elevation of 8600 m above sea level (the Himalayas and Karakoram mountain ranges) to a minimum of 0 m (Arabian Sea).

Pakistan has a diverse climate with four distinct climate regions, i.e., glacial, humid, arid, and hyper-arid regions. Extreme north of the country is occupied by glacial region with the mean elevation of 4158 m and mean annual precipitation of 348 mm. This region is mostly covered with snow and glaciers. The Hindu Kush Himalaya (HKH) mountain ranges, which are very famous for snow and glaciers, also lie in the glacial region. Snow and glacier melt from this region in the summer season is the main source of water in the Indus river for domestic, industrial, and agriculture sectors of Pakistan. However, the excessive snow and glacier melt caused acute flood events in Pakistan, especially the 2010 flood, which took thousands of lives and severely damaged the infrastructure and economy of the country. The humid region consists of very high mountains of HKH and all major rivers of the country (Indus, Kabul, Gilgit, Chitral, Swat, Panjkora, Hunza, Kurram, and Jhelum). Mean elevation of the humid region is 1286 m, with a mean annual precipitation of 852 mm. The arid region is composed of major agricultural region of Punjab province. The Indus river and its tributaries flow through the

region and considered as a primary source of water in the region. The average elevation and mean annual precipitation in arid region are 633 m and 322 mm, respectively. Hyper-arid region includes the Sindh and Balochistan provinces, and south part of the Punjab province. Most of the hyper-arid region consists of barren lands and dry mountains. The mean elevation and mean annual precipitation in the region are 444 m and 133 mm, respectively.



Figure 1. (a) Elevation of Pakistan (study area) obtained from Shuttle Radar Topography Model (SRTM), (b) Geographic location of rain gauges and four climate regions, and (c) Mean annual precipitation over Pakistan during 2000–2015.

2.2. Ground Precipitation Gauge Data

Ground precipitation gauges (GPGs) are considered as standard precipitation data sources, which provide direct precipitation records for satellite precipitation calibration and validation processes. Pakistan Meteorology Department (PMD), and Snow and Ice Hydrology Project (SIHP) of Water and Power Development Authority (WAPDA) own the meteorological data of Pakistan. SIHP operates the meteorological stations at high altitudes, mostly situated in glacial and humid regions. The daily precipitation records from 2000 to 2015 at 102 GPGs have been collected from both organizations. Out of 102 GPGs, precipitation records of 79 GPGs are collected from PMD, while the remaining 23 GPGs from WAPDA. Figure 1b shows the geographical distribution of obtained meteorological stations. In the current study, GPGs are named with respect to the climate zones, i.e., GPGs in the glacial, humid, arid, and hyper-arid regions are represented as G-GPGs, H-GPGs, A-GPGs, and HA-GPGs. The area (km²)/number of stations in glacial, humid, arid, and hyper-arid regions are 72,774/19, 137,753/39, 270,484/19, and 322,929/25, respectively.

The precipitation records are manually collected by PMD and WAPDA, which might be subjected to personal and instrumental errors. Furthermore, GPGs located at high elevations may also be subjected to other errors due to splashing and wind. Therefore, PMD and WAPDA do the correction of precipitation records following the standard code of World Meteorological Organization (WMO-N). Moreover, data quality tests, Skewness and Kurtosis methods are performed after filling the missing data using the zero-order method [16].

The distribution of GPGs in Pakistan is non-homogenous and very limited/scarce, which is not enough to understand the precipitation pattern, its spatial and temporal variability over different climate regions [48]. In the last few decades, the number of GPGs has significantly increased, but the density is still not able to cope with scientific requirements [49]. Therefore, the application of SPPs and development of MSPDs are of paramount importance over Pakistan as an alternative precipitation source for hydrological and meteorological applications.

2.3. Satellite-based Precipitation Datasets

The experiment is conducted using three satellite-based precipitation datasets, including TMPA 3B42-v7, CMORPH, and PERSIANN-CDR, and one re-analysis precipitation dataset (Era-Interim). These datasets are selected because they are comprehensively evaluated in the previous studies over the study area (Pakistan).

2.3.1. TMPA 3B42-v7

The Tropical Rainfall Measurement Mission (TRMM), a first dedicated satellite precipitation product launched in late 1997, was capable of measuring moderate and heavy precipitation with reasonable accuracy. The TRMM Multi-Satellite Precipitation Analysis (TMPA) is one of the most suitable TRMM based multi-satellite precipitation products for near real-time (3B42-RT) and research (3B42-V6/3B42-V7) applications [50,51]. To enhance the calibration process, 3B42-V7 uses the global real-time precipitation datasets from Global Precipitation Climatology Centre (GPCC). In the current study, TMPA 3B42-V7 (with spatial resolution of 0.25° and daily temporal resolution) is used which has many advanced specifications as compared to previous versions such as a newer IR brightness temperature dataset, latitude band calibration system, additional satellite inputs, and a single uniformed processed surface precipitation analysis [31].

2.3.2. PERSIANN-CDR

Center of Hydrometeorology and Remote Sensing (CHRS) at the University of California Irvine developed the PERSIANN algorithm and has 0.25°/daily spatial/temporal resolution. PERSIANN adjust the neural network parameters by using Passive Microwave (PMW) data (TMI, AMSU-B, and SSM/I) to enhance the accuracy of precipitation estimates. PERSIANN-CDR, developed by CHRS with the same spatial resolution and family, uses the same neural network used by previous products for precipitation estimation. The only difference is the use of input IR dataset with GridSat-B1 instead of CPC-IR. Moreover, PMW data is not used in PERSIANN-CDR [52].

2.3.3. CMORPH

CMORPH, developed by Climate Prediction Center (CPC), uses IR based motion vectors derived from PMW data to propagate high-quality precipitation [53]. CMORPH has 0.25° spatial and 3

hours temporal resolution. IR data integrate the observations from GOES 8, GOES 10, Meteosat 5, Meteosat 8, and GMS 5 satellites. Defense Meteorological Satellite Program (DMSP), TRMM, and NOAA polar-orbiting operational meteorological satellites are used in PMW derived precipitation estimation in CMORPH.

2.3.4. Reanalysis Precipitation Product (Era-Interim)

The global atmospheric reanalysis precipitation dataset, Era-Interim, is developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). Era-interim provides the real-time global precipitation observations from 1979 to present with spatial resolution 0.25°. Era-Interim data is generated from the data assimilation system based on IFS (Cy31r2) 2006 release. It includes a four-dimensional variational (4D-Var) analysis with 12 hours window. Era-Interim estimates the precipitation based on temperature and relative humidity while using weather forecast model [54].

2.4. Dynamic Bayesian Model Averaging (DBMA)

The term dynamic is used to consider both spatial and temporal variability of BMA weights. BMA is a statistical approach to combine the predictions and inferences based on multiple competing models to yield a highly skillful and reliable probabilistic merging [26,27]. BMA uses Bayes theorem to model not only parameters uncertainty through prior distribution but also model the uncertainty by obtaining parameter and model's posterior. It is a robust technique for direct model selection, combined estimation, and prediction. BMA estimates weights as posterior probabilities, which reflect model's forecasting skills relative to other models in the training period [27]. In the current study, BMA is applied to optimally merge four SPPs to improve satellite precipitation estimation and obtain best agreement with GPGs measurements by adjusting predictive probability density function (PDF). Figure 2 shows the flow chart of dynamic BMA (DBMA) algorithm for merging multi-satellite precipitation. MSPD in the current study is developed by following the methodology proposed by Ma et al. [26], to which readers are referred for a detailed description.

The calculated BMA weights are adjusted for all merged members at each calibrated location on each day. Note that BMA analyses are performed only when all the SPPs captured precipitation (hit cases only). Since BMA weights are applied at the point scale, pixels having at least a single GPG are checked for spatiotemporal SPPs coincidence. The next day weights of dynamic BMA are based on adding new sample data (e.g., 41st day) and deleting the oldest sample data (e.g., 1st day) from the training dataset. Therefore, there is considerable overlap (39 days) between the training data at GPG from one day to another.

Several studies demonstrated that the performance of dynamic MSPDs in terms of Root Mean Square Error (RMSE) and Correlation Coefficient (CC) increased till 40 days; however, no improvements are observed beyond 40 days interval [26,47,55–57]. Therefore, BMA is trained for a time window of three years using 40 days training period. The training was performed for each calibrated GPG on a daily temporal scale. Considering year 2002 as a basic (starting) year (for which previous two years of daily precipitation data is available), 40 days of training period in one year and 40 training days in each of the previous two years, with a total of 120 days are considered to optimize BMA weights. Hamill [55] also concluded that expanding the training period to include the same season days from previous year would likely improve model performance. The optimal BMA weights were then interpolated to the whole study area (Pakistan) using the ordinary kriging method. Finally, a blended satellite precipitation estimate for each grid is calculated based on individual data and corresponding optimal grid weights across Pakistan.

To check the importance and sensitivity of training time windows (TTW), we have considered four sets of TTW, i.e., 1-year, 2-years, 3-years, and 5-years. MAE and RMSE statistical indices calculated on daily temporal scale averaged over each year of TTW are used to evaluate the accuracy and performance of MSPD during each TTW (Figures 3 and 4, respectively). Figures 3 and 4 show the MSPD performance in glacial, humid, arid, and hyper-arid regions during four sets of TTW. The analyses show that performance of DBMA-MSPD increases with an increase in TTW duration, which supports

the statement of Hamil [55]. Higher fluctuations (low accuracy) are observed for 1-year TTW, which are decreasing with an increase in TTW duration, e.g., 3-years and 5-years. The analyses depict that there are no significant improvements when we expand the duration from 3-years to 5-years (only minor improvements in 5-years TTW as compared to 3-years). Therefore, 3-years of TTW is selected because of considerable improvements as compared to 1-year and 2-years TTW and less number of training datasets as compared to 5-years TTW.



Figure 2. The flow chart of Dynamic Bayesian Model Averaging (DBMA).



Figure 3. Performance evaluation of DBMA-MSPD using averaged (annual) MAE during 1-year, 2-years, 3-years, and 5-years TWW in (**a**) Glacial, (**b**) Humid, (**c**) Arid, and (**d**) Hyper-arid regions.



Figure 4. Performance evaluation of DBMA-MSPD using averaged (annual) RMSE during 1-year, 2-years, 3-years, and 5-years TWW in (**a**) Glacial (**b**) Humid (**c**) Arid, and (**d**) Hyper-arid regions.

2.5. Performance Evaluation

The statistical indices, including Mean Bias Error (MBE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Correlation Coefficient (CC) [58], Kling-Gupta efficiency (KGE score) [59], Theil's U coefficient [60], ensemble spread score (ESS), and standard deviation (SD) are used to assess and compare the performance of DBMA-MSPD against the GPG observations. Table 1 represents a complete description of the statistical indices. MBE demonstrates over/under-estimation of precipitation with positive/negative values. MAE represents the average absolute error between DBMA-MSPD and GPGs. RMSE evaluates the average squared error magnitude between the merged (DBMA-MSPD) and observed precipitation data. CC represents an agreement between the simulated and observed observations. KGE combines the variability ratio (γ), bias ratio (β), and correlation coefficient (R). The perfect values for KGE, β , and γ are all 1. Theil's U evaluates the accuracy of DBMA-MSPD forecasted precipitation associated with GPGs. Theil's U can be related to R2 but differs in boundary conditions, i.e., not bounded by zero and one. Value of Theil's U closer to zero represents perfect forecasting accuracy; assumes a value of 1 when models forecast the same errors as the naïve no-change extrapolation and value greater than 1 represent worst forecasting accuracy and has to be rejected [60]. ESS is usually the deviation of DBMA-MSPD precipitation from its mean [61]. The ideal DBMA merged precipitation will have the same size of ESS as its RMSE at the same time span [61,62]. SD is used to describe the MSPD simulated precipitation variance. A higher SD value indicates a higher variation in simulated precipitation from its mean.

Moreover, we further compared the performance of DBMA-MSPD with SMA based MSPD. SMA is a fixed weight ensemble approach, which merges SPPs using the arithmetic mean of all members [41]. The equation for SMA approach is given as:

$$MSPD_{SMA} = \frac{1}{n} \sum_{i=1}^{n} S_i$$
(1)

where S_i represents individual SPP, n is the number of SPPs and MSPD_{SMA} is the final MSPD developed using SMA approach.

Table 1. Statistical indices used to evaluate the performance of DBMA-MSPD. <i>M</i> is the simulated
(merged precipitation) data, O is the observed precipitation data from GPGs, n is the number of samples,
X is the data element ($X = M$ for the DBMA-MSPD while $X = O$ for GPGs), CV is the coefficient of
variation, and bars on variables represents the average values.

Statistical Index	Equation	Perfect Value
Mean Bias Error (MBE)	$\text{MBE} = \frac{1}{n} \sum_{i=1}^{n} (M_i - O_i)$	0
Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^{n} M_i - O_i $	0
Root Mean Square Error (RMSE)	$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (M_i - O_i)^2}$	0
Correlation Coefficient (CC)	$CC = \frac{\sum\limits_{i=1}^n (M_i - \overline{M_i}) (O_i - \overline{O_i})}{\sqrt{\sum\limits_{i=1}^n (M_i - \overline{M_i})^2} \sqrt{\sum\limits_{i=1}^n (O_i - \overline{O_i})^2}}$	1
KGE Score	KGE = $1 - \sqrt{(CC - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$ where $\beta = \overline{M}/\overline{O} \ \gamma = (CV)_{12}/(CV)_{22}$	1
Theil's U	$U = \sqrt{\sum_{i=1}^{n} (M_i - O_i)^2 / \sum_{i=1}^{n} M_i^2}$	0
Ensemble spread score	$ESS = \sqrt{rac{1}{n-1}\sum\limits_{i=1}^{n} \left(\overline{M} - M_i ight)^2}$	0
Standard deviation	$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (X - \overline{X})^2}$	

The performance of DBMA-MSPD is compared with fixed weight MSPD (SMA, hereinafter SMA-MSPD) using a skill score for each statistical index used in the analyses (shown in Table 1). For example, the skill scores of MBE and MAE are shown in Equations (2) and (3), respectively.

$$MBE_{score} = 100 \times \left(1 - \frac{MBE_{DBMA}}{MBE_{SMA}}\right)$$
(2)

$$MAE_{score} = 100 \times \left(1 - \frac{MAE_{DBMA}}{MAE_{SMA}}\right)$$
(3)

3. Results

3.1. Data Normality and Box-Cox Transformation

BMA analyses require the normal distribution of precipitation data; however, the precipitation data is usually gamma-distributed. Since we have considered only the hit cases and 2 mm as a threshold for precipitation and no precipitation events, the data is closely normally distributed almost overall GPGs/SPPs in four climate regions. The Jarque-Bera test was used to determine the normality of precipitation distribution [63]. Data is normally distributed when the skewness (symmetry of residuals) is close to zero, and kurtosis (peakedness of the curve) is closer to a value of 3. During our analyses, skewness of the data is ranging between 0.27 to 0.78, kurtosis between 2.18 to 3.99, and Jarque-Bera between 13.46 to 47.28. Besides the Jarque-Bera test, we also applied Box-Cox transformation before the application of BMA to ensure that the transformed individual is close to variables [26,42,44,64]. Box-Cox transformation has the capability to transform the non-normally distributed data to datasets either normally distributed or close enough to the normal distribution [65]. The results confirm that precipitation is normally distributed after Box-Cox transformation (Figure 5).



Figure 5. Distribution of precipitation of dominant ensemble member (TMPA) over representative GPG across (**a**) Glacial, (**b**) Humid, (**c**) Arid, and (**d**) Hyper-arid regions. Blue dots show the distribution of probability density function (PDF) along the normal (red) line.

The application of Box-Cox transformation is dependent on λ values, which might be positive or negative. This demonstrates that the transformed results do not cover the entire range $(-\infty, +\infty)$; therefore, approximately normal distribution of precipitation is expected. The λ is obtained by calculating a minimum of unconstrained multivariable function using the Nelder-Mead simplex technique [66]. The λ value is dependent on length of the sampling data (number of precipitation observations), and its values for TMPA, PERSIANN-CDR, Era-Interim, and CMORPH are ranging between 0.02–1.57, -0.03–2.21, 0.02–1.96, and -0.08–2.37, respectively. The λ values of TMPA at representative GPGs across each glacial, humid, arid, and hyper-arid regions are 1.352, 0.515, 0.619, and 1.386, respectively.

3.2. Spatiotemporal Distribution of DBMA-MSPD Weights over Pakistan

The weights of the merged members (TMPA, Era-Interim, PERSIANN-CDR, and CMORPH) for the DBMA-MSPD are averaged over four seasons during 2000–2015 and presented in Figure 6. The analysis shows that DBMA assigned higher weights to TMPA (Figure 6a) and PERSIANN-CDR (Figure 6c), followed by Era-Interim (Figure 6b) and CMOPRH (Figure 6d). The average weights of TMPA, PERSIANN-CDR, Era-Interim, and CMORPH during 2000–2015 are 0.32, 0.27, 0.22, and 0.19, respectively. On the regional scale, TMPA shows higher skills in glacial (0.32) and humid (0.37) regions as compared to PERSIANN-CDR having DBMA weights of 0.27 and 0.25. Moreover, the average weights of Era-Interim vary from a maximum of 0.22 in glacial to 0.20 in humid region. Similarly, CMORPH weights

varies from 0.18 (humid region) to 0.20 (glacial region). However, in arid region PERSIANN-CDR (0.29) dominated TMPA (0.26), followed by Era-Interim (0.25). Further, hyper-arid region is dominated by TMPA and Era-Interim with the average weights of 0.31 and 0.29, respectively. While the average weights of PERSIANN-CDR and CMORPH in hyper-arid regions are 0.22 and 0.18.



Figure 6. Spatial distribution of DBMA-MSPD average weights of merged members, (**a**) TMPA, (**b**) Era-Interim, (**c**) PERSIANN-CDR, and (**d**) CMORPH, during 2000–2015.

The seasonal (pre-monsoon, monsoon, post-monsoon, and winter) distribution of DBMA-MSPD weights of four merged members over Pakistan during 2000–2015 are shown in Figures 7–10. The temporal spans of four seasons are: April, May, and June (pre-monsoon), July, August, and September (monsoon), October and November (post-monsoon), and December, January, February, and March (winter). More than 60% of precipitation is received during the monsoon season in Pakistan, which varies with topography and climate in magnitude from low (<100 mm) in the glacial region



Figure 7. Spatial distribution of DBMA-MSPD weights of merged members, (**a**) TMPA, (**b**) Era-Interim, (**c**) PERSIANN-CDR, and (**d**) CMORPH, in pre-monsoon season during 2000–2015.



Figure 8. Spatial distribution of DBMA-MSPD weights of merged members, (**a**) TMPA, (**b**) Era-Interim, (**c**) PERSIANN-CDR, and (**d**) CMORPH, in monsoon season during 2000–2015.

During the pre-monsoon season (Figure 7), higher weights are assigned to TMPA (0.31), followed by Era-Interim (0.25). Weights of PERSIANN-CDR and CMORPH are 0.23 and 0.21. Highest average weight for TMPA is observed in the arid (0.36) and humid (0.33) regions. The distribution of weight is gradually decreasing from arid towards the hyper-arid region (0.31). However, for Era-Interim, highest average weight is observed in the humid (0.34) and glacial (0.25) regions. In contrast, for PERSIANN-CDR, the highest weight is observed in hyper-arid (0.28) and arid regions (0.26). Moreover, no specific trend is observed for CMORPH.



Figure 9. Spatial distribution of DBMA-MSPD weights of merged members, (**a**) TMPA, (**b**) Era-Interim, (**c**) PERSIANN-CDR, and (**d**) CMORPH, in post-monsoon season during 2000–2015.

During the monsoon season (Figure 8), TMPA dominates other merging members having higher accuracy with an overall average weight of 0.52, which is significantly higher than average weights of Era-Interim (0.21), PERSIANN-CDR (0.14) and CMORPH (0.13), respectively. Similarly, on a regional scale, TMPA shows relatively higher skills in the humid and hyper-arid regions with average weights of 0.54 and 0.53, respectively. TMPA experienced almost similar average weights in all the climate regions. In the case of Era-Interim, highest weights are observed in hyper-arid and arid regions with average values of 0.24 and 0.23, respectively. However, no particular trend is observed for PERSIANN-CDR and CMORPH. PERSIANN-CDR and CMORPH experience higher skills in the glacial region with average weights of 0.16 and 0.17, respectively.



Figure 10. Spatial distribution of relative DBMA-MSPD weights of the merged members, (**a**) TMPA, (**b**) Era-Interim, (**c**) PERSIANN-CDR, and (**d**) CMORPH, in winter season during 2000–2015.

During the post-monsoon season (Figure 9), weights are shifted to PERSIANN-CDR and CMORPH with average weights of 0.33 and 0.25. The average weights of TMPA and Era-Interim are 0.20 and 0.22. PERSIANN-CDR and CMORPH show a relatively small variation in the distribution of weights across different climate regions. PERSIANN-CDR shows higher skills in arid (0.37) and hyper-arid (0.36) regions while relatively lower skills in glacial (0.33) and humid (0.32) regions. However, no significant variation is observed for CMORPH. The regional average weights for CMORPH are 0.26 (glacial), 0.27 (humid), 0.26 (arid) and 0.25 (hyper-arid).

During the winter season (Figure 10), PERSIANN-CDR and TMPA show higher skills as compared to other merged members. The average weights of the merged members during the winter season are 0.33 (TMPA), 0.14 (Era-Interim), 0.38 (PERSIANN-CDR), and 0.15 (CMORPH). In the glacial, arid and hyper-arid regions, PERSIANN-CDR dominated TMPA. However, TMPA shows slightly higher skills in the humid region. The average weights of PERSIANN-CDR in all the climate regions from glacial to

hyper-arid region are 0.35, 0.38, 0.41, and 0.40, respectively. Similarly, the average weight of TMPA are 0.27 (glacial), 0.39 (humid), 0.28 (arid) and 0.35 (hyper-arid).

The average DBMA-MSPD weights (Figure 11) at representative GPGs in each climate region are plotted against DOY (Day of the Year). Representative GPGs are those whose weights are close to the average temporal distribution of DBMA-MSPD weights in the corresponding climate region. The representative stations for glacial, humid, arid and hyper-arid regions are G-GPG10, H-GPG17, A-GPG13, and HA-GPG17, respectively. In the glacial region, significant variation is observed in PERSIANN-CDR and Era-Interim weights. However, less volatility can be seen for TMPA and moderate for CMORPH. Peak weights of TMPA, PERSIANN-CDR, and Era-Interim are observed between DOY 156-232 during monsoon season (late pre-monsoon, monsoon, and early post-monsoon rainy days). In the humid region, higher weights are assigned to TMPA followed by PERSIANN-CDR. Higher fluctuations in weights are observed only for PERSIANN-CDR. The monsoon precipitation (DOY 181-273) is captured very well by TMPA and PERSIANN-CDR representing peak weights during the season. In contrast, less volatility is observed for Era-Interim and CMORPH. The arid region is dominated by PERSIANN-CDR, followed by TMPA and Era-Interim (almost similar weights). There is a significant fluctuation of PERSIANN-CDR and Era-Interim weights in the monsoon season. TMPA is showing a different trend in the arid region as compared to other regions, i.e., no significant fluctuation during the monsoon season. In the hyper-arid region, dominated by TMPA and Era-Interim, all the SPPs captured high precipitation events effectively.



Figure 11. Distribution of DBMA-MSPD weights against the DOY at representative GPGs in (**a**) Glacial, (**b**) Humid, (**c**) Arid, and (**d**) Hyper-arid regions.

3.3. Statistical Evaluation of DBMA-MSPD over Pakistan

Statistical evaluation of DBMA-MSPD is performed using pixel by pixel analysis (pixels having at least one GPG were selected) to ensure proper analysis of DBMA-MSPD over Pakistan. The DBMA-MSPD performance in daily precipitation estimation over 102 GPGs during the past 16 years over entire Pakistan is shown in Figure 12.



Figure 12. Spatial distribution of DBMA-MSPD daily statistical indices: (**a**) MBE, (**b**) MAE, (**c**) RMSE, (**d**) CC, (**e**) KGE score, and (**f**) Theil's U over Pakistan.

Figure 12a shows mean bias error (MBE) of DBMA-MSPD over Pakistan. DBMA-MSPD overestimated precipitation in the extreme north (glacial region) of Pakistan; however, the error

is declining towards the south. The maximum overestimation is observed at HMS4 (+1.89 mm/day), while the maximum underestimation is -1.14 mm/day at HMS34.

Figure 12b shows the spatial distribution of mean absolute error (MAE) over Pakistan. A high magnitude of error is observed in the glacial and humid regions, which are gradually decreasing toward the hyper-arid region. The MAE for DBMA-MSPD is variable even in the glacial and humid regions ranging from 2.69 mm/day (at HMS4) to 1.36 mm/day at GMS5. However, MAE falls gradually from humid region to hyper-arid region ranging from 2.38 mm/day at HMS11 to 0.71 mm/day at HAMS18.

Figure 12c shows RMSE distribution, where a similar spatial distribution trend as MAE is observed, i.e., RMSE is decreasing from north to south. RMSE is ranging from 11.96 mm/day (at HMS11) to 1.72 mm/day (HAMS8). The figure depicts higher RMSE in the downstream part of the glacial region and upstream part of the humid region. Minimum RMSE is observed in the South-West of Pakistan.

DBMA-MSPD shows a higher correlation with the GPGs with an average CC of 0.70 (Figure 12d) over the entire Pakistan. Higher CC is mainly observed in the south-east of Pakistan (maximum CC is 0.84 at HAMS18), while lower CC is 0.41 observed at GMS2 (glacial region). DBMA-MSPD shows poor correlation with GPGs in the glacial region, which gradually increasing towards the South of Pakistan. The average CC in the glacial region is 0.55, which increases to 0.68 in humid region, 0.77 in arid, and 0.81 in hyper-arid regions.

Figure 12e shows the spatial distribution of KGE score compared to GPGs observations. Smaller KGE score is observed in extreme east and west of glacial region, which gradually increases toward the south of Pakistan (from glacial region to hyper-arid region). The maximum KGE score is observed in hyper-arid region, more specifically at HA-GPG20 (0.93) and HA-GPG19 (0.91). Smallest KGE score in glacial the region might be associated with poor performance of individual SPPs in glacial and humid regions. The smallest KGE score is 0.33 at G-GPG10 and 0.34 at G-GPG11.

Figure 12f represents the spatial distribution of Theil's U over Pakistan. Theil's U is used in this study to observe the merged dataset forecasting accuracy over the spatial span of study area. It also squares the deviation to give more weight to errors and to exaggerate errors that help to detect the regions with high errors. Theil's U values closer to or larger than 1 represent poor forecasting capability of the statistical model (DBMA). It is observed that the model's forecasting is considerably increasing from north to south of Pakistan. The performance of DBMA-MSPD is poor in the glacial region, which is increasing gradually towards the hyper-arid region. Theil's U is ranging from 0.70 at H-GPG4 to 0.37 at HA-GPG18. The average Theil's U overall climate regions are 0.53, representing better forecasting accuracy of DBMA-MSPD.

Table 2 shows the mean and median of ESS and SD scores of DBMA-MSPD and SMA-MSPD during different seasons across four climate regions of Pakistan. For DBMA-MSPD, the analyses show that high ensemble spread is observed during monsoon season followed by pre- and post-monsoon seasons. The minimum spread is depicted in hyper-arid region across all seasons due to low precipitation magnitude/intensity and higher SPPs performance. On the other hand, higher spread is observed in humid region (more specifically in the north-west of the region) with maximum average ESS values of 11.38 mm/day and 9.26 mm/day, respectively, during monsoon and pre-monsoon seasons. A similar trend is also observed for the glacial region, with the maximum average ESS values of 10.26 mm/day in monsoon, and 8.80 mm/day in pre-monsoon seasons. In contrast, higher ESS values for arid region are observed during the post-monsoon season, which might be associated with relatively frequent precipitation (moderate magnitude and intensity).

Results in Table 2 show that variation in precipitation is dependent on magnitude and intensity, i.e., higher variation during monsoon season and lower variation in winter season. SD statistic for DBMA-MSPD shows high variations in glacial region as compared with other regions. Lower variations are observed in hyper-arid region, having very minimal precipitation amount. The maximum mean SD values in glacial and humid regions are 12.58 mm/day and 11.43 mm/day during monsoon season.

Season	Climate Region	ESS Mean		ESS Median		SD Mean		SD Median	
Scason		DBMA	SMA	DBMA	SMA	DBMA	SMA	DBMA	SMA
Pre-Monsoon	Glacial	8.80	4.86	8.71	4.84	11.35	6.56	11.44	6.50
	Humid	9.26	6.60	8.94	6.49	10.42	6.25	10.47	6.23
	Arid	3.97	2.58	3.69	2.67	4.21	2.73	3.98	2.69
	Hyper-arid	3.03	1.90	2.99	1.94	3.63	2.25	3.55	2.33
Monsoon	Glacial	10.26	5.33	9.61	5.36	12.58	6.79	12.47	6.75
	Humid	11.38	6.67	11.25	6.60	11.43	6.57	11.58	6.63
	Arid	3.76	2.27	3.50	2.24	5.95	3.92	5.74	3.95
	Hyper-arid	2.67	1.81	1.92	1.85	4.48	2.91	4.41	2.98
Post-Monsoon	Glacial	9.41	5.12	9.20	5.10	10.80	6.33	10.92	6.41
	Humid	8.14	4.67	7.86	4.56	9.24	5.29	9.46	5.28
	Arid	4.14	2.58	3.95	2.60	4.35	2.83	4.10	2.88
	Hyper-arid	3.13	1.95	3.02	2.02	3.61	2.33	3.53	2.38
Winter	Glacial	8.79	4.87	8.75	4.75	10.61	6.07	10.70	6.03
	Humid	8.67	4.73	8.43	4.80	8.69	4.93	8.46	4.92
	Arid	3.99	2.48	3.71	2.51	4.19	2.68	3.95	2.72
	Hyper-arid	2.75	1.77	2.88	1.80	3.45	2.12	3.53	2.10

Table 2. Ensemble Spread Scores (ESS) and Standard deviations (SD) of DBMA-MSPD and SMA-MSPD across four climate regions during four seasons (mm/day).

SMA-MSPD shows comparatively lower ensemble spread and precipitation variation as compared to DBMA-MSPD. The analyses show that ESS score is increasing with the increase in precipitation magnitude and intensity. Thus, a higher ESS score is observed during monsoon followed by post-monsoon and pre-monsoon seasons. The ensemble spread of SMA-MSPD also varies with topography, i.e., high ESS score (6.67 mm/day) across humid region followed by 5.33 mm/day in glacial region during monsoon season. Minimum ensemble spread is observed across hyper-arid region, which might be associated with minimal precipitation in the region. A minimum ESS score of 1.77 mm/day is observed during the winter season across hyper-arid region. The analyses show that SMA-MSPD is not as effective as DBMA-MSPD to express the ensemble spread in precipitation during all seasons across all climate regions.

SMA-MSPD has a relatively poor capacity to depict the variation in precipitation as DBMA-MSPD evaluated across all four climate regions during different precipitation seasons. The precipitation variation is dependent on seasonal variations in magnitude and intensity of precipitation. In glacial region, a higher SD (average value of 6.79 mm/day) is observed during monsoon season, while minimum SD in 6.07 mm/day in winter season. The higher variation in SMA-MSPD precipitation (with average SD value of 6.57 mm/day) in humid region is observed during the monsoon season, while minimum average SD (2.12 mm/day) is observed during winter season in hyper-arid region.

Figure 13 shows the skill scores (improvements) of DBMA-MSPD compared with SMA-MSPD across Pakistan. Figure 13a shows higher improvements in MBE in the center north of glacial region (45–50%), which is decreasing from north to south (from glacial to humid, arid, and hyper-arid regions). Higher skills in glacial and humid regions confirm the significant improvements of BMA algorithm as compared to SMA. Besides, the supremacy dynamic (spatial and temporal) variation of weights can also be depicted from glacial and humid regions. Lower improvements in MBE are observed in hyper-arid region, more specifically in the south-east. The lower skills in hyper-arid region are associated with higher performance of individual SPP in the region.



Figure 13. Spatial variation of skill scores, (**a**) MBE, (**b**) MAE, (**c**) RMSE, (**d**) CC, (**e**) KGE, and (**f**) Theil's U, of DBMA-MSPD compared with SMA-MSPD.

MAE skills show higher improvements in 40%–50% and 30–40% in glacial and humid regions, respectively (Figure 13b). Slightly lower improvements in MAE are observed in the west of arid and east of hyper-arid regions ranging between 30–35%. Figure 13c shows higher improvements in RMSE across glacial (40–50%) and humid (40–45%) region, while lower in the middle of hyper-arid region (30–35%).

Lower improvements as compared to other skill scores are observed in CC, KGE, and Theil's U (Figure 13d–f). DBMA-MSPD significantly improved the CC in east of glacial region as compared to SMA-MSPD, while moderate improvements are observed in the rest of glacial and humid regions. Lower CC skill score is depicted in the extreme west of hyper-arid region (10–15%). There is no particular trend observed in the KGE skill score in glacial and humid regions; however, higher improvements are observed in these two regions, which are declining towards the hyper-arid region (Figure 13e). Theil's U score has been significantly increased in the extreme north of the glacial region (22–24%). However, the improvements of DBMA-MSPD show a decreasing trend towards the south (hyper-arid region) of Pakistan.

The ESS and SD skill scores of DBMA-MSPD compared with SMA-MSPD are also calculated and presented in Figure 14. Higher improvements in ensemble spread are observed in humid (40–50%) and glacial (35–40%) regions. ESS score is decreasing towards the hyper-arid region, where the ESS score is ranging between 20–30% (Figure 14a). SD skill score (Figure 14b) shows contrasting figues as compared with ESS score in humid and glacial regions. Maximum SD skill score (35–45%) is observed in glacial region, while SD score is ranging from 30% to 45% in humid region. Minimum improvements (20–30%) in terms of SD are observed in hyper-arid region.



Figure 14. Spatial distribution of (a) ESS and (b) SD skill scores of DBMA-MSPD compared with SMA-MSPD.

Overall, the skill scores show high improvements of DBMA-MSPD in glacial and humid regions, which is declining towards the hyper-arid region. Lower improvements in hyper-arid/arid region are due to the high performance of individual SPPs, low magnitude of precipitation and low elevation.

Higher improvements in glacial region are due to the robustness of BMA algorithm and dynamic variation of weights as compared to fixed weight SMA-MSPD.

3.4. Comparison of DBMA-MSPD with SMA-MSPD and Merging Members using GPGs Observations

The comparison and evaluation of DBMA-MSPD and four merged members against 102 GPGs in four climate zones using selected six statistical indices are presented in Table 3. The analysis shows that DBMA-MSPD outperformed all the merged members and showing significant improvement in all climate zones. Among the merged members, TMPA shows supremacy in capturing precipitation with high forecasting accuracies and a high correlation with GPGs in all climate zones. On the other side, CMORPH presents the worst performance in all climate zones. The analysis confirms a significant reduction in errors for DBMA-MSPD as compared to individual merged members and shows highest potential to be utilized in hydrological applications [26,48].

Zone	SPPs/MSPD	MBE (mm/day)	MAE (mm/day)	RMSE (mm/day)	CC	KGE Score	Theil's U
	DBMA-MSPD	0.99	1.96	8.80	0.55	0.36	0.53
	SMA-MSPD	1.85	2.66	10.18	0.48	0.32	0.55
Glacial	TMPA	1.91	2.79	10.83	0.45	0.31	0.56
zone	Era-Interim	2.27	3.25	11.39	0.38	0.28	0.64
	PERSIANN-CDR	2.14	3.01	11.01	0.42	0.27	0.60
	CMORPH	2.51	3.60	11.87	0.34	0.25	0.67
Humid zone	DBMA-MSPD	0.78	1.75	8.23	0.68	0.51	0.44
	SMA-MSPD	1.47	2.64	10.99	0.61	0.47	0.46
	TMPA	1.54	2.80	11.62	0.59	0.46	0.48
	Era-Interim	1.86	3.34	11.58	0.52	0.43	0.54
	PERSIANN-CDR	1.65	3.05	11.32	0.56	0.40	0.51
	CMORPH	2.04	3.65	11.84	0.49	0.38	0.58
Arid	DBMA-MSPD	-0.25	1.52	4.62	0.78	0.67	0.41
	SMA-MSPD	-0.32	2.87	6.79	0.73	0.63	0.44
	TMPA	-0.36	3.15	7.28	0.71	0.62	0.46
zone	Era-Interim	-0.53	3.65	7.91	0.63	0.58	0.54
	PERSIANN-CDR	0.46	3.40	7.64	0.67	0.55	0.50
	CMORPH	-0.62	3.91	8.18	0.61	0.53	0.57
Hyper-Arid zone	DBMA-MSPD	-0.71	1.23	3.83	0.80	0.86	0.41
	SMA-MSPD	-1.43	2.16	5.72	0.76	0.81	0.44
	TMPA	-1.51	2.32	5.97	0.76	0.80	0.46
	Era-Interim	-2.00	2.82	6.53	0.68	0.77	0.53
	PERSIANN-CDR	-1.76	2.57	6.24	0.72	0.75	0.50
	CMORPH	-2.25	3.04	6.84	0.65	0.71	0.57

Table 3. Evaluation of DBMA-MSPD and four merged members against 102 GPGs using daily averaged statistical indices during 2000–2015.

The improvements of DBMA-MSPD against SMA-MSPD are 46.48%, 26.31%, 13.55%, 12.73%, 11.11%, and 3.63% in MBE, MAE, RMSE, CC, KGE and Theil's U, respectively, across glacial region. Similarly, 46.94%, 21.87% and 50.34% in MBE, 33.71%, 47.04% and 43.05% in MAE, 25.11%, 31.96% and 33.04% in RMSE, 10.29%, 6.41% and 5% in CC, 7.84%, 5.97% and 5.81 in KGE score, and 4.35%, 6.82% and 6.82% in Theil's U improvements are observed in humid, arid and hyper-arid regions, respectively.

DBMA-MSPD is compared with TMPA (best merging member) to quantify improvement of developed MSPD. DBMA-MSPD improved MBE, MAE, RMSE, CC, KGE score and Theil's U in the glacial zone by 48.17%, 29.75%,18.74%, 18.18%, 13.88% and 5.36%, respectively. Performance of DBMA-MSPD in humid zone is increased by 49.35% (MBE), 37.50% (MAE), 29.17% (RMSE), 13.23% (CC), 9.80% (KGE score) and 8.33% (Theil's U). Similarly, improvements in arid and hyper-arid zones are 30.55% and 52.98% in MBE, 51.74% and 46.98% in MAE, 36.54% and 35.84% in RMSE, 8.97% and

5% in CC, 7.46% and 6.97% in KGE score, respectively. Equal improvements of 10.87% in Theil's U are observed in both arid and hyper-arid zones. The average improvements of MSPD across all four climate regions are 45.26% (MBE), 30.99% (MAE), 30.1% (RMSE), 11.34% (CC), 9.53% (KGE score) and 8.86% (Theil's U).

4. Discussion

Accurate estimation and capturing higher spatial and temporal variation in precipitation is a challenging task over the diverse climate and complex topography of Pakistan [16]. The sparse distribution of GPGs contributes to low accuracy in estimation of precise precipitation. Therefore, SPPs are used as an alternative which provides homogenous precipitation estimates on regional to global scales. With significant advancement in remote sensing techniques and continuous improvements in satellite-based retrieval algorithms provide cost-effective and reliable precipitation estimates [68,69]. However, SPPs observations are subjected to uncertainties resulting from retrieval algorithms. To reduce these uncertainties and to obtain high-quality precipitation estimates, researchers have focused on merging individual SPPs [16,22,26,70].

Very limited studies are found which focused on the development of merged precipitation datasets over Pakistan. Very recently, Rahman et al. [48], Rahman et al. [16], and Muhammad et al. [22] developed and evaluated merged multi-satellite precipitation datasets (MSPDs) in Pakistan. Muhammad et al. [22] developed MSPD across Pakistan by assigning relative weights to SPPs, i.e., Global Precipitation Measurement (GPM)-based Integrated Multi-Satellite Retrievals for GPM (IMERG) research version (IMERG-IR), IMERG real-time (IT), and TRMM 2B42 (RT). The results demonstrated significant agreement with GPG observations and improvements as compared to individual SPPs. Rahman et al. [48] developed Dynamic Clustered Bayesian Model Averaging (DCBA-MSPD) across Pakistan and evaluated its performance across four climate regions (spatial evaluation) and four seasons (temporal evaluation) during 2000–2015. They concluded better performance of DCBA-MSPD as compared to individual SPP members. However, relatively poor performance was observed over high elevated areas of glacial and humid regions. Similarly, Rahman et al. [16] reported significant improvement in precipitation estimation and reduction in uncertainties both on regional and seasonal scales. In the current study, the performance of DBMA-MSPD shows elevation dependency. Besides the significant improvements in reduction of errors associated with individual SPPs, higher discrepancies are observed in glacial and humid zones (high elevated zones). Besides the improvements, all the MSPDs show a strong dependency on precipitation magnitude/intensity and elevation [16,26,41]. Shen et al. [41] reported that precipitation estimation capability of MSPDs reduces significantly at an elevation higher than 4000 m.

The possible sources of errors are elevation dependency, climate variability, topographic complexity, impact (limitations) of different sensors used in SPPs, and the retrieval algorithms used to estimate precipitation [48,71]. The topography of Pakistan is very complex, consisting of the permanent snow and glaciers in glacial region, high mountain peaks and hilly areas in northern parts and plain areas having arid and hyper-arid characteristics in the south, all in a very short spatial span. The topographic complexity also contributes to strong scattering of MW (microwave) signals, especially across cold climate and snow-dominated regions (i.e., glacial region) [31,72].

The poor performance of all MSPDs evaluated over Pakistan in the glacial region might be associated with the performance of IR (InfraRed) and MW sensors used in precipitation estimation. Information about precipitation estimates based on the lowest temperature on the top of cloud is provided by IR sensors, while PMW provides details about precipitation area rather than cloud temperature [73]. The clouds over glacial region are relatively warm and hamper the capabilities of sensors to estimate precipitation from warmer clouds. This is because the cloud top temperature is too warm for the IR sensors' threshold to differentiate between precipitation and no precipitation clouds [74,75]. Furthermore, the accuracy of IR sensors is also dependent on brightness temperature, and polarization properties vary with snow cover and exposure, which in turn depend on altitude and

terrain of the mountainous region [72,76]. Besides the limitations of IR sensors, PMW sensors have poor accuracy in detecting precipitation over mountains (glacial and humid regions in the current study), which is due to warmer clouds that could produce heavy precipitation without more ice aloft in PMW algorithm, dense vegetation cover, and relative coarse spatial resolution [73].

Intense precipitation during monsoon season across the humid region causes attenuation in signals, which is very frequent in such regions [77]. In addition, the retrieval algorithms used to estimate precipitation also contributes to errors in SPPs [78,79]. The algorithms used in SPPs uses IR brightness temperature at the top of the cloud and indirectly estimate the precipitation without taking into consideration the elevation impact and sub-cloud evaporation phenomena [72,73], hence significantly affect the quality and accuracy of SPPs [80]. Besides the poor performance of SPPs due to retrieval algorithms over heavy precipitation regions, the processing schemes of MW and IR sensors contribute to further uncertainties [72]. Moreover, external errors related to the data quality of GPGs such as splashing and wind effects, human-induced errors, impact of snow, and non-uniform/sparse/limited distribution of GPGs also plays a role in calibrating the SPPs [81].

The factors mentioned above are the causes of poor performance of MSPDs over glacial and humid regions of Pakistan, i.e., poor performance of SPPs in glacial and humid regions leads to greater uncertainties in the developed MSPDs [57]. High performance of individual SPPs in arid and hyper-arid regions [16,82,83] contributes to the high performance of MSPDs.

MSPD developed using a dynamic algorithm (variation of weights over time and space) has many advantages as compared to fixed weight-based MSPD. The dynamic weights can account the regional topographic complexity as we all as extreme local climate that can influence the merging process. In comparison to fixed weight MSPDs, the improvements of DBMA-MSPD is quantified using a skill score calculated using SMA-MSPD as a reference forecast. The analyses show significant improvement of DBMA-MSPD across glacial and humid regions and relatively lower improvements in hyper-arid region. We investigated the DBMA-MSPD performance for a particular season at a specific location. Variation of weights over time in a dynamic algorithm demonstrates the performance of MSPDs based on precipitation intensity and magnitude, e.g., monsoon season in Pakistan having an intense and high magnitude of precipitation. Sometimes, the dynamic algorithms (such as DBMA and DCBA) randomly select an individual SPP rather than considering the SPP having a high correlation with GPGs and low errors (high weight) in a particular season. Thus, higher uncertainties are introduced during the merging process [56].

5. Conclusions

The current study proposed a merged multi-satellite precipitation dataset (MSPD) using Dynamic Bayesian Model Averaging (DBMA) algorithm from four satellite-based precipitation products (SPPs), including TMPA 3B42v7, PERSIANN-CDR, Era-Interim, and CMORPH. The MSPD was evaluated over four different climate regions of Pakistan, i.e., glacial, humid, arid and hyper-arid, both on spatial and seasonal (pre-monsoon, monsoon, post-monsoon, and winter) scale. Performance of MSPD is evaluated over 102 GPGs using six statistical indices at daily temporal scale, including Mean Bias Error (MBE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Correlation Coefficient (CC), Kling-Gupta efficiency (KGE score) and Theil's U coefficient during 2000-2015. The key findings from the current study are listed below:

- The MSPD has significantly improved the performance of individual SPPs. TMPA has higher accuracy as compared to other SPPs. The average improvements of MSPD across all climate regions with respect to TMPA are 45.26% (MBE), 30.99% (MAE), 30.1% (RMSE), 11.34% (CC), 9.53% (KGE score) and 8.86% (Theil's U).
- (2) DBMA-MSPD has assigned higher weights to TMPA and PERSIANN-CDR followed by Era-Interim and CMORPH. The average weights of TMPA, PERSIANN-CDR, Era-Interim, and CMORPH across Pakistan during 2000–2015 are 0.32, 0.27, 0.22 and 0.19, respectively.

- (3) On regional scale, TMPA shows higher skills in glacial (0.32) and humid (0.37) regions as compared to PERSIANN-CDR with 0.27 (glacial) and 0.25 (humid), Era-Interim with 0.22 (glacial) and 0.20 (humid), and CMORPH with 0.20 (glacial) and 0.18 (humid). Moreover, arid and hyper-arid regions are dominated by PERSIANN-CDR and TMPA, with average weights of 0.29 and 0.31, respectively.
- (4) Spatial evaluation of MSPD depicted poorer performance in glacial and humid regions, which significantly improved towards arid and hyper-arid regions. Precipitation is overestimated in glacial and humid regions while underestimated in arid and hyper-arid regions. Maximum overestimation and underestimation are +1.89 mm/day and -1.14 mm/day, respectively. MAE and RMSE are ranging from 2.69–0.71 mm/day and 11.96–1.72 mm/day, respectively. Higher CC is observed in hyper-arid (0.84) while lower in glacial (0.41). The average CC across glacial, humid, arid, and hyper-arid regions are 0.55, 0.68, 0.77, and 0.81, respectively. The maximum (0.93) and minim (0.33) KGE score is observed in hyper-arid and glacial regions, indicating better performance of MSPD in hyper-arid region. Theil's U is ranging from 0.70 (humid) to 0.37 (hyper-arid) with an average value of 0.53 across all climate regions.
- (5) The heavy precipitation seasons (pre-monsoon and monsoon) are dominated by TMPA (with average weights of 0.31 and 0.52) and Era-Interim (0.25 and 0.21). PERSIANN-CDR presented a higher performance in post-monsoon and winter seasons with average weights of 0.33 and 0.38, respectively. The significant seasonal variations of DBMA-MSPD weights indicate the necessity to use dynamic weights in merging the SPPs.
- (6) TMPA and PERSIANN-CDR accurately captured the precipitation trends in heavy precipitation seasons (pre-monsoon and monsoon) in glacial and humid regions. Similarly, in arid and hyper-arid regions PERSIANN-CDR and Era-Interim and TMPA and Era-Interim followed the heavy precipitation trends.
- (7) The ensemble spread and variation of DBMA-MSPD precipitation are evaluated using ensemble spread score (ESS) and standard deviation (SD), respectively. The ensemble spread is increasing with magnitude and intensity of the precipitation. Higher ESS with an average value of 11.38 mm/day was observed across the humid region during monsoon season. Lower ensemble spread is depicted in hyper-arid region during the winter season. The variation in precipitation showed elevation dependency, subjected to magnitude and intensity of precipitation, i.e., higher variation during monsoon season over high elevation regions. The maximum average SD values in glacial and humid regions are 12.58 mm/day and 11.43 mm/day.
- (8) The ESS and SD scores of SMA-MSPD showed relatively lower ensemble spread and variation in precipitation. Maximum spread with average value of 6.67 mm/day is observed during monsoon season in humid region, while a lower ESS score of 1.77 mm/day is observed in hyper-arid region during the winter season. Similarly, higher variation in SMA-MSPD precipitation (with SD average value of 6.79 mm/day) is depicted in glacial region during monsoon season, while minimum SD (2.12 mm/day) in hyper-arid region during winter season.
- (9) The skill score used to quantify the improvements of DBMA-MSPD against SMA-MSPD shows high improvements (40–50% for MBE, MAE and RMSE, and 20–25% for CC, KGE and Theil's U) in glacial and humid region. However, relatively low improvements are observed across hyper-arid region (20–30% for MBE, MAE, and RMSE, and 10–15% for CC, KGE, and Theil's U). Higher improvements in ensemble spread are observed in humid (40–50%) and glacial (35–40%) regions, while lower improvements in hyper-arid region (20–30%). Similarly, maximum and minimum improvements are observed in SD across glacial (35–45%) and hyper-arid (20–30%) regions.

In further studies, the role and impact of different SPPs on the performance of MSPD can be explored, e.g., by replacing the Era-Interim with its new version ERA5, the Era-Interim with SM2RAIN-based products, the TMPA with IMERG. Moreover, real-time versions of TMPA and IMERG can also be considered.

Author Contributions: Methodology of the paper was designed by K.U.R. and S.S., K.U.R. and Y.W. analyzed the data, K.U.R. and S.S. interpreted the results, K.U.R. wrote the paper, Y.W. and M.S. provided the suggestions and reviews. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the National Natural Science Foundation of China (Grant numbers 51779119 and 51839006).

Acknowledgments: The authors gratefully acknowledge the Pakistan Meteorology Department (PMD) and Water and Power Development Authority (WAPDA) for providing GPGs data overall Pakistan. The authors present their gratitude to SPPs developers. Authors greatly appreciate the efforts of anonymous reviewers, whose comments and suggestions significantly improved the quality of the current manuscript.

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

Abbreviations

SPPs	Satellite precipitation products
MSPD	Merged multi-Satellite Precipitation Dataset
DBMA	Dynamic Bayesian Model Averaging
DBMA-MSPD	DBMA-Merged multi-Satellite Precipitation Dataset
GPGs	Ground Precipitation Gauges
SMA	Simple Model Averaging
SMA-MSPD	Simple Model Averaging-Merged multi-Satellite Precipitation Dataset
	Dynamic Clustered Bayesian Model Averaging-Merged multi-Satellite
DCDA-INI5PD	Precipitation Dataset

References

- 1. Kidd, C.; Levizzani, V.; Laviola, S. Quantitative precipitation estimation from Earth observation satellites. *Rainfall State Sci.* **2010**, *191*, 127–158.
- 2. 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]
- 3. Huffman, G.J.; Adler, R.F.; Morrissey, M.M.; Bolvin, D.T.; Curtis, S.; Joyce, R.; McGavock, B.; Susskind, J. Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.* **2001**, *2*, 36–50. [CrossRef]
- 4. Zhang, Z.; Tian, J.; Huang, Y.; Chen, X.; Chen, S.; Duan, Z. Hydrologic evaluation of TRMM and GPM IMERG satellite-based precipitation in a humid basin of China. *Remote Sens.* **2019**, *11*, 431. [CrossRef]
- 5. 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]
- 6. 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]
- 7. 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]
- Zhu, Q.; Xuan, W.; Liu, L.; Xu, Y.P. Evaluation and hydrological application of precipitation estimates derived from PERSIANN-CDR, TRMM 3B42V7, and NCEP-CFSR over humid regions in China. *Hydrol. Process.* 2016, *30*, 3061–3083. [CrossRef]
- 9. Belabid, N.; Zhao, F.; Brocca, L.; Huang, Y.; Tan, Y. Near-real-time flood forecasting based on satellite precipitation products. *Remote Sens.* **2019**, *11*, 252. [CrossRef]
- 10. Yuan, F.; Zhang, L.; Soe, K.M.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Applications of TRMM-and GPM-era multiple-satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in myanmar. *Remote Sens.* **2019**, *11*, 140. [CrossRef]
- 11. Tao, H.; Fischer, T.; Zeng, Y.; Fraedrich, K. Evaluation of TRMM 3B43 precipitation data for drought monitoring in Jiangsu Province, China. *Water* **2016**, *8*, 221. [CrossRef]

- Wang, F.; Yang, H.; Wang, Z.; Zhang, Z.; Li, Z. Drought evaluation with CMORPH satellite precipitation data in the Yellow River Basin by using gridded standardized precipitation evapotranspiration index. *Remote Sens.* 2019, 11, 485. [CrossRef]
- 13. Zhang, X.; Su, Z.; Lv, J.; Liu, W.; Ma, M.; Peng, J.; Leng, G. A set of satellite-based near real-time meteorological drought monitoring data over China. *Remote Sens.* **2019**, *11*, 453. [CrossRef]
- 14. Gehne, M.; Hamill, T.M.; Kiladis, G.N.; Trenberth, K.E. Comparison of global precipitation estimates across a range of temporal and spatial scales. *J. Clim.* **2016**, *29*, 7773–7795. [CrossRef]
- 15. Gebregiorgis, A.S.; Hossain, F. How well can we estimate error variance of satellite precipitation data around the world? *Atmos. Res.* **2015**, *154*, 39–59. [CrossRef]
- 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]
- 17. 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]
- Yi, L.; 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. [CrossRef]
- 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]
- 20. 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]
- 21. 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]
- 22. 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]
- 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. Duan, Q.; Phillips, T.J. Bayesian estimation of local signal and noise in multimodel simulations of climate change. *J. Geophys. Res. Atmos.* **2010**, *115*, D18123. [CrossRef]
- 25. 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]
- Ma, Y.; Hong, Y.; Chen, Y.; Yang, Y.; Tang, G.; Yao, Y.; Long, D.; Li, C.; Han, Z.; Liu, R. Performance of optimally merged multisatellite precipitation products using the dynamic Bayesian model averaging scheme over the Tibetan Plateau. *J. Geophys. Res. Atmos.* 2018, *123*, 814–834. [CrossRef]
- 27. Raftery, A.E.; Gneiting, T.; Balabdaoui, F.; Polakowski, M. Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* 2005, *133*, 1155–1174. [CrossRef]
- 28. Pereira Filho, A.J. Integrating gauge, radar and satellite rainfall. In Proceedings of the WWRP International Precipitation Working Group Workshop, CGMS-WMO, Monterey, CA, USA, 25–28 October 2004.
- 29. Sinclair, S.; Pegram, G. Combining radar and rain gauge rainfall estimates using conditional merging. *Atmos. Sci. Lett.* **2005**, *6*, 19–22. [CrossRef]
- Vila, D.A.; De Goncalves, L.G.G.; Toll, D.L.; Rozante, J.R. Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. *J. Hydrometeorol.* 2009, 10, 533–543. [CrossRef]
- 31. 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]
- 32. Hill, D.J. Assimilation of weather radar and binary ubiquitous sensor measurements for quantitative precipitation estimation. *J. Hydroinformatics* **2015**, *17*, 598–613. [CrossRef]

- 33. Xie, P.; Xiong, A.Y. A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. *J. Geophys. Res. Atmos.* **2011**, *116*, D21106. [CrossRef]
- 34. Li, H.; Hong, Y.; Xie, P.; Gao, J.; Niu, Z.; Kirstetter, P.; Yong, B. Variational merged of hourly gauge-satellite precipitation in China: Preliminary results. *J. Geophys. Res. Atmos.* **2015**, *120*, 9897–9915. [CrossRef]
- 35. Long, Y.; Zhang, Y.; Ma, Q. A merging framework for rainfall estimation at high spatiotemporal resolution for distributed hydrological modeling in a data-scarce area. *Remote Sens.* **2016**, *8*, 599. [CrossRef]
- 36. Li, M.; Shao, Q. An improved statistical approach to merge satellite rainfall estimates and raingauge data. *J. Hydrol.* **2010**, *385*, 51–64. [CrossRef]
- 37. Ochoa-Rodriguez, S.; Wang, L.P.; Willems, P.; Onof, C. A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resour. Res.* **2019**, *55*, 6356–6391. [CrossRef]
- Sivasubramaniam, K.; Sharma, A.; Alfredsen, K. Merging radar and gauge information within a dynamical model combination framework for precipitation estimation in cold climates. *Environ. Model. Softw.* 2019, 119, 99–110. [CrossRef]
- Yang, Z.; Hsu, K.; Sorooshian, S.; Xu, X.; Braithwaite, D.; Zhang, Y.; Verbist, K.M. Merging high-resolution satellite-based precipitation fields and point-scale rain gauge measurements—A case study in Chile. *J. Geophys. Res. Atmos.* 2017, 122, 5267–5284. [CrossRef]
- 40. Wu, H.; Zhang, X.; Liang, S.; Yang, H.; Zhou, G. Estimation of clear-sky land surface longwave radiation from MODIS data products by merging multiple models. *J. Geophys. Res. Atmos.* **2012**, *117*, D22107. [CrossRef]
- 41. Shen, Y.; Xiong, A.; Hong, Y.; Yu, J.; Pan, Y.; Chen, Z.; Saharia, M. Uncertainty analysis of five satellite-based precipitation products and evaluation of three optimally merged multi-algorithm products over the Tibetan Plateau. *Int. J. Remote Sens.* **2014**, *35*, 6843–6858. [CrossRef]
- 42. Chen, Y.; Yuan, W.; Xia, J.; Fisher, J.B.; Dong, W.; Zhang, X.; Liang, S.; Ye, A.; Cai, W.; Feng, J. Using Bayesian model averaging to estimate terrestrial evapotranspiration in China. *J. Hydrol.* **2015**, *528*, 537–549. [CrossRef]
- Yao, Y.; Liang, S.; Li, X.; Hong, Y.; Fisher, J.B.; Zhang, N.; Chen, J.; Cheng, J.; Zhao, S.; Zhang, X. Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. J. Geophys. Res. Atmos. 2014, 119, 4521–4545. [CrossRef]
- 44. Duan, Q.; Ajami, N.K.; Gao, X.; Sorooshian, S. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv. Water Resour.* **2007**, *30*, 1371–1386. [CrossRef]
- 45. Jiang, S.; Ren, L.; Xu, C.-Y.; Liu, S.; Yuan, F.; Yang, X. Quantifying multi-source uncertainties in multi-model predictions using the Bayesian model averaging scheme. *Hydrol. Res.* **2018**, *49*, 954–970. [CrossRef]
- 46. Roy, T.; Serrat-Capdevila, A.; Gupta, H.; Valdes, J. A platform for probabilistic multimodel and multiproduct streamflow forecasting. *Water Resour. Res.* **2017**, *53*, 376–399. [CrossRef]
- Liu, R.; Ma, Y.; Yang, Y.; Han, Z.; Tang, G.; Liu, Q.; Hong, Y. Error analysis of ensemble multi-satellite precipitation datasets over the Tibetan Plateau. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 4684–4687.
- Rahman, K.U.; Shang, S.; Shahid, M.; Wen, Y.; Khan, Z. Application of Dynamic Clustered Bayesian Model Averaging (DCBA) algorithm for merging multi-satellite precipitation products over Pakistan. *J. Hydrometeorol.* 2019. [CrossRef]
- 49. Rahman, K.U.; Shang, S.; Shahid, M.; Wen, Y. Performance assessment of SM2RAIN-CCI and SM2RAIN-ASCAT precipitation products over Pakistan. *Remote Sens.* **2019**, *11*, 2040. [CrossRef]
- 50. Liu, Z. Comparison of precipitation estimates between Version 7 3-hourly TRMM Multi-Satellite Precipitation Analysis (TMPA) near-real-time and research products. *Atmos. Res.* **2015**, *153*, 119–133. [CrossRef]
- Prakash, S.; Mitra, A.K.; AghaKouchak, A.; Pai, D. Error characterization of TRMM Multisatellite Precipitation Analysis (TMPA-3B42) products over India for different seasons. J. Hydrol. 2015, 529, 1302–1312. [CrossRef]
- 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]
- 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]
- 54. Dee, D.P.; Uppala, S.; Simmons, A.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.; Balsamo, G.; Bauer, D.P. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [CrossRef]

- 55. Hamill, T.M. Comments on "Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging". *Mon. Weather Rev.* **2007**, *135*, 4226–4230. [CrossRef]
- Wilson, L.J.; Beauregard, S.; Raftery, A.E.; Verret, R. Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *Mon. Weather Rev.* 2007, 135, 1364–1385. [CrossRef]
- 57. Sloughter, J.M.L.; Raftery, A.E.; Gneiting, T.; Fraley, C. Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Mon. Weather Rev.* **2007**, *135*, 3209–3220. [CrossRef]
- 58. Ebert, E.E. Methods for verifying satellite precipitation estimates. In *Measuring Precipitation from Space;* Springer: Berlin/Heidelberg, Germany, 2007; pp. 345–356.
- 59. Kling, H.; Fuchs, M.; Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* **2012**, *424*, 264–277. [CrossRef]
- 60. Bliemel, F. *Theil's Forecast Accuracy Coefficient: A Clarification;* SAGE Publications Sage CA: Los Angeles, CA, USA, 1973.
- 61. Zhu, Y. Ensemble forecast: A new approach to uncertainty and predictability. *Adv. Atmos. Sci.* **2005**, *22*, 781–788. [CrossRef]
- 62. Buizza, R.; Houtekamer, P.; Pellerin, G.; Toth, Z.; Zhu, Y.; Wei, M. A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Weather Rev.* **2005**, *133*, 1076–1097. [CrossRef]
- 63. Jarque, C.M.; Bera, A.K. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ. Lett.* **1980**, *6*, 255–259. [CrossRef]
- 64. Fu, Y.; Xia, J.; Yuan, W.; Xu, B.; Wu, X.; Chen, Y.; Zhang, H. Assessment of multiple precipitation products over major river basins of China. *Theor. Appl. Climatol.* **2016**, *123*, 11–22. [CrossRef]
- 65. Renard, B. A Bayesian hierarchical approach to regional frequency analysis. *Water Resour. Res.* **2011**, *47*, W11513. [CrossRef]
- 66. Lagarias, J.C.; Reeds, J.A.; Wright, M.H.; Wright, P.E. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optim.* **1998**, *9*, 112–147. [CrossRef]
- Khan, S.I.; Hong, Y.; Gourley, J.J.; Khattak, M.U.K.; Yong, B.; Vergara, H.J. Evaluation of three high-resolution satellite precipitation estimates: Potential for monsoon monitoring over Pakistan. *Adv. Space Res.* 2014, 54, 670–684. [CrossRef]
- 68. Hobouchian, M.P.; Salio, P.; Skabar, Y.G.; Vila, D.; Garreaud, R. Assessment of satellite precipitation estimates over the slopes of the subtropical Andes. *Atmos. Res.* **2017**, *190*, 43–54. [CrossRef]
- Tapiador, F.J.; Turk, F.J.; Petersen, W.; Hou, A.Y.; García-Ortega, E.; Machado, L.A.; Angelis, C.F.; Salio, P.; Kidd, C.; Huffman, G.J. Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.* 2012, 104, 70–97. [CrossRef]
- Massari, C.; Camici, S.; Ciabatta, L.; Brocca, L. Exploiting satellite-based surface soil moisture for flood forecasting in the Mediterranean area: State update versus rainfall correction. *Remote Sens.* 2018, 10, 292. [CrossRef]
- Beighley, R.E.; Ray, R.; He, Y.; Lee, H.; Schaller, L.; Andreadis, K.; Durand, M.; Alsdorf, D.; Shum, C. Comparing satellite derived precipitation datasets using the Hillslope River Routing (HRR) model in the Congo River Basin. *Hydrol. Process.* 2011, 25, 3216–3229. [CrossRef]
- 72. Scheel, M.; Rohrer, M.; Huggel, C.; Santos Villar, D.; Silvestre, E.; Huffman, G. 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]
- 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*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 193–204.
- 74. Bitew, M.M.; Gebremichael, M. Evaluation through independent measurements: Complex terrain and humid tropical region in Ethiopia. In *Satellite Rainfall Applications for Surface Hydrology*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 205–214.
- 75. Hong, Y.; Gochis, D.; Cheng, J.-T.; Hsu, K.-L.; Sorooshian, S. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *J. Hydrometeorol.* **2007**, *8*, 469–482. [CrossRef]
- 76. Amlien, J. Remote sensing of snow with passive microwave radiometers—A review of current algorithms. *Report* **2008**, *1019*, 52.

- 77. Villarini, G.; Krajewski, W.F. Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surv. Geophys.* **2010**, *31*, 107–129. [CrossRef]
- 78. AghaKouchak, A.; Nasrollahi, N.; Habib, E. Accounting for uncertainties of the TRMM satellite estimates. *Remote Sens.* **2009**, *1*, 606–619. [CrossRef]
- 79. Yan, J.; Gebremichael, M. Estimating actual rainfall from satellite rainfall products. *Atmos. Res.* **2009**, *92*, 481–488. [CrossRef]
- 80. Li, X.; Zhang, Q.; Xu, C.-Y. Assessing the performance of satellite-based precipitation products and its dependence on topography over Poyang Lake basin. *Theor. Appl. Climatol.* **2014**, *115*, 713–729. [CrossRef]
- 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]
- 82. Hussain, Y.; Satgé, F.; Hussain, M.B.; Martinez-Carvajal, H.; Bonnet, M.-P.; Cárdenas-Soto, M.; Roig, H.L.; Akhter, G. Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theor. Appl. Climatol.* **2018**, *131*, 1119–1132. [CrossRef]
- 83. Iqbal, M.F.; Athar, H. Validation of satellite based precipitation over diverse topography of Pakistan. *Atmos. Res.* **2018**, 201, 247–260. [CrossRef]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).