



# Article Evaluation of the Performance of Three Satellite Precipitation Products over Africa

# Aleix Serrat-Capdevila<sup>1,2,\*</sup>, Manuel Merino<sup>3</sup>, Juan B Valdes<sup>2</sup> and Matej Durcik<sup>4</sup>

- <sup>1</sup> Water Global Practice, The World Bank, Washington, DC 20009, USA
- <sup>2</sup> Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, AZ 85721, USA; jvaldes@email.arizona.edu
- <sup>3</sup> Montgomery and Associates Ltd., Santiago 7550120, Chile; manuel.merino@gmail.com
- <sup>4</sup> Biosphere 2, The University of Arizona, Tucson, AZ 85721, USA; mdurcik@email.arizona.edu
- \* Correspondence: aserratcapdevila@worldbank.org; Tel.: +1-520-979-6438

Academic Editors: Zhongbo Su, Yijian Zeng, Zoltan Vekerdy, Magaly Koch, Richard Müller and Prasad S. Thenkabail Received: 19 May 2016; Accepted: 22 September 2016; Published: 13 October 2016

Abstract: We present an evaluation of daily estimates from three near real-time quasi-global Satellite Precipitation Products—Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), and Climate Prediction Center (CPC) Morphing Technique (CMORPH)—over the African continent, using the Global Precipitation Climatology Project one Degree Day (GPCP-1dd) as a reference dataset for years 2001 to 2013. Different types of errors are characterized for each season as a function of spatial classifications (latitudinal bands, climatic zones and topography) and in relationship with the main rain-producing mechanisms in the continent: the Intertropical Convergence Zone (ITCZ) and the East African Monsoon. A bias correction of the satellite estimates is applied using a probability density function (pdf) matching approach, with a bias analysis as a function of rain intensity, season and latitude. The effects of bias correction on different error terms are analyzed, showing an almost elimination of the mean and variance terms in most of the cases. While raw estimates of TMPA show higher efficiency, all products have similar efficiencies after bias correction. PERSIANN consistently shows the smallest median errors when it correctly detects precipitation events. The areas with smallest relative errors and other performance measures follow the position of the ITCZ oscillating seasonally over the equator, illustrating the close relationship between satellite estimates and rainfall regime.

**Keywords:** satellite precipitation products; Africa; ITCZ; error analysis; bias correction; PERSIANN; CMORPH; TMPA 3B42-RT; product averaging

# 1. Introduction

Characterizing precipitation in space and time is still a challenge. As the main input component of the water balance, rainfall is a key control of hydrologic variables such as evapotranspiration, soil moisture, streamflow and others, and thus essential to accurate estimation. While ground gauges can provide good measurements of point precipitation at the ground level, the number of functional and recording stations has progressively and significantly decreased since the 1980s [1,2].

For the last few decades, satellites and remote sensing technology have been providing numerous spatially distributed but indirect estimates of precipitation. These are indirect because they are derived from measurements of other parameters: cloud top temperatures and reflectivity using infrared (IR), passive micro-waves (PMW) or radar (active micro waves, R) which are used to estimate rainfall in or below the cloud—likely to be different from rainfall reaching the ground level. Thus, as the parameters

2 of 22

of the intensity–reflectivity curve (Z-R relation) can change from storm to storm, these estimates come with errors that vary depending on the type of storm systems, geography, topography and cloud characteristics. Satellite Precipitation Products (SPPs) merge available data from sensors in different satellites to provide homogeneous gridded precipitation fields at a specific resolution and time aggregation. Errors depend on the number and type of sensors taking measurements over a specific location at a given time, and the algorithms used to assimilate the available data into a coherent gridded dataset.

The evaluation of the performance of SPPs and the characterization of their errors have been the object of numerous publications. Perhaps due to the difficulty of having good reference datasets for evaluation, the focus of most studies has often been local or regional in Africa [3–15], with few continental or global evaluation efforts [16,17]. This is also illustrated at a global level by Maggioni et al. (2016), providing a review of SPP evaluation studies of Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) Real Time (RT) and Version 7 (V7), Global Satellite Mapping of Precipitation (GSMaP), Climate Prediction Center (CPC) Morphing Technique (CMORPH), and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) with and without Cloud Classification System (CCS), across six continents [18]. To address the lack of global or continent-wide analysis and further the understanding of the performance of SPPs over Africa, this paper presents an evaluation of three near real-time quasi-global satellite precipitation products over the African continent, compared against a quasi-global reference dataset. With an emphasis on the near real-time aspect of the products, and thus their potential for operational monitoring and predictive use, the goal of this work is to characterize their errors in relationship to latitude, topography, climate and season. Specifically, we analyze how errors vary with the dynamics of the main climatic mechanisms controlling rainfall in Africa such as the ITCZ. We also explore what error components can be reduced through bias correction techniques, as well as how and where model averaging approaches produce better estimates. The three quasi-global and near real-time products evaluated here are CMORPH, TMPA 3B42-RT, and PERSIANN. We have chosen the Global Precipitation Climatology Program-one degree daily (GPCP-1dd) as our reference dataset for being a well-established climatology dataset integrating ground observations and a range of satellite precipitation estimates over the continent. The results presented here illustrate the need for bias correction for any use of SPPs, and can inform their hydrologic applications in Africa, such as real-time monitoring of rainfall over poorly gauged regions and for streamflow monitoring and prediction (especially in medium to large basins where concentration times are high).

#### 2. Background

Numerous case studies evaluate the performance of SPPs over specific locations, comparing them against local datasets including ground data and radar observations when available. In perhaps the most spatially extensive validation synthesis, Ebert et al. (2007) reports validation efforts covering the US, Australia and Northwestern Europe, finding that performance of SPP estimates is highly dependent on the rainfall regime, being more accurate in low latitudes and in summer, especially when the precipitation regime is controlled by deep convection [16]. When estimating extreme events, performance of SPPs has been reported to decrease significantly in terms of biases and ability to correctly detect extreme rainfall [19–21]. Different elevation and aspect dependent biases have been observed for SPPs in complex topography in various case studies. In general, performance tends to be poorer over mountains but locally calibrated products may do better [11,22–25]. Detailed reviews on the accuracy of SPPs, past evaluation efforts and their applications, can be found in Demaria and Serrat-Capdevila (2015) and Serrat-Capdevila et al. (2013) [26,27].

In Africa, Adeyewa and Nakamura (2003) evaluate monthly and seasonal precipitation over main climate zones and find that the TRMM Precipitation Radar, an IR rainfall estimate (TMPI) and a monthly bias corrected TMPA product tended to overestimate rainfall in the tropical zone for the northern hemisphere winter and spring and produced better estimates during the rest of the year. High biases were found during dry seasons in semi-arid areas [28].

In the Great Rift Valley in Ethiopia, PERSIANN and CMORPH overestimated the average monthly rainfall by 43% and 11%, respectively, and TMPA underestimated it by 5%, compared to ground gauge measurements. In the tropical region, CMORPH and TMPA 3B42-RT overestimated rainfall at low elevation, but gave reasonable results at higher elevations, while PERSIANN results were better in the low lands but degraded with altitude. In semi-arid climate, all the products underestimated rainfall [29].

Streamflow simulations in the Nzoia Basin in Kenya, using TMPA real time estimates with a 0.5 correlation with the basin rain gauge, had a bias of 20% and a Nash-Sutcliffe (N-S) coefficient of 0.57. The use of the bias corrected research quality product TMPA 3B42-V6 improved the bias to 3.6% and the N-S efficiency to 0.71 [30].

Thiemig et al. (2012) evaluated the performance of six SPP—CMORPH, Rainfall Estimation Algorithm Version 2 (RFE2.0), TRMM 3B42, Goddard Profiling Algorithm (GPROF), PERSIANN and GSMap Moving Vector with Kalman filter (MVK)—over the Zambezi, Volta, Baro-Akobo and Juba Shabelle basins on a daily, monthly and annual basis [14]. The performance of SPPs was evaluated against a topography-based interpolation of rain gauge data, and was generally superior in the tropical zones. Poorer performance was found in semiarid, mountainous regions, and in capturing events on the extremes of the distribution. Over the basins in tropical zones, PERSIANN and CMORPH overestimate both the number of rainy days and precipitation rates. However, while PERSIANN underestimates rainfall in mountainous areas due to problems detecting high altitude and light rain, CMORPH has a superior ability to capture rain better in mountainous areas. The TRMM 3B42 showed the best performance, capturing seasonal variability, seasonal patterns, average annual precipitation and peaks. This is likely due to its incorporated bias correction from ground gauge networks information in the research product.

On the other hand, while Thiemig et al. (2012) provide an excellent analysis over three study areas in Africa using rain-gauge data, few studies exist that compare SPPs over large regions of the globe such as the African Continent, and that relate a spatially continuous analysis with the main climatic mechanisms that control precipitation dynamics. This is partly due to the challenges of obtaining an adequate reference dataset with a large regional coverage and with consistent quality. Relating the performance and error distributions of SPPs with continental scale precipitation-generating climatic mechanisms is especially difficult without a consistent "ground-truth" accurately characterizing precipitation in space and time.

The locations with high-quality reference datasets including dense rain gauge networks and/or stage IV radar products assimilating available ground gauges are very scarce and hardly exist in developing regions. In addition, catchment level rainfall errors from the use of a few rain gauges can be comparable to errors of satellite estimates aggregated over the catchment [31]. Even in data-rich regions, an accurate definition of the precipitation "ground truth" is difficult.

This difficulty has led other research teams to adopt other approaches, such as comparing satellite precipitation products between each other instead of against a global reference dataset. Tian and Peters-Lidard (2010) computed a map of global measurement uncertainties as the spread of a six satellite dataset ensemble, finding smaller deviations from the average (40%–60%) over the oceans, especially on the tropics, and over South America, and the largest uncertainties (100%–140%) on latitudes over 40°, and mostly over the cold season [17]. They also detected higher uncertainties over complex terrain, such as the Tibetan Plateau, the Rockies, the Andes, as well as coastlines and water bodies. The estimates agree more over tropical, convection driven rainfall regimes, and less on frontal and orographic precipitation, as found by many others. Thus, such dependence of the uncertainty and disagreement among SPPs on geography, topography and type of storms seems to be consistent with regional performance as found by Ebert et al. (2007). The lack of high-quality dense ground-gauge networks, even at the regional level, lead Casse et al. (2015) to use the 3B42-V7 research

product as a reference dataset against which to compare and bias-correct other products, over the choice of a gauge-only product (CPC) in their Niger sub-basin study area. The gauge-only product was found to smooth the rainfall, spreading it over more rainy days and spatial area, due to the low density of available gauges, highlighting the advantages of multi-sensor products integrating different measurements [32]. On the other hand, Ringard et al. (2015) found in their SPP evaluations over French Guiana and North Brazil that 3B42-V7 was only better in areas of intense precipitation and especially with extreme precipitation events. In drier areas all SPP showed similar performance, with PERSIANN equaling that of 3B42-V7; and in regions along the Amazon, PERSIANN was best [33].

The efforts presented here contribute to the understanding of the spatial and temporal dynamics of near real-time satellite precipitation errors over the African continent by building upon the efforts described above, such as global quantifications of SPP uncertainty, regional validations, and basin-specific comparisons in Africa against ground gauge data. In this paper, we present a comparison of three near real-time SPPs against a reference dataset over the African continent through the quantification of different error metrics and their spatial distributions across latitude, topography, climate and season. Given that basis, we examine the spatial nature of biases, the efficiency of bias correction and evaluate potential improvements of averaging different SPPs. The choice of evaluating near real-time products (and not their a posteriori gauge-adjusted, improved research versions) responds to their operational usefulness for real-time rainfall estimation in large poorly-gauged areas and their use in real-time streamflow monitoring and prediction, especially in large basins with high concentration times.

One of the main challenges in the evaluation of satellite precipitation products is the choice of the reference "ground-truth" dataset. Precipitation can be highly heterogeneous in space and time and is thus difficult to monitor and represent accurately. Ground measurements and satellite estimates of precipitation are observations of different nature and are difficult to compare and integrate. Rain-gauge measurements represent the rainfall in a few square centimeters over intervals of seconds to days and with a spatial characterization limited by the number of rain-gauges. In developing settings, these are rarely available in real time. Satellite rainfall estimates are indirect measures (from infrared, passive microwave or radar sensors) and often have spatial resolutions ranging from 0.04° and 0.25°, where a precipitation value is representative of a 16 km<sup>2</sup> and 625 km<sup>2</sup>, respectively, and time steps ranging from half hour to three hours or a day. While satellite estimates provide a spatially distributed picture with average precipitation values over their grid-cells, rain-gauges provide accurate direct measurements at the point scale. It is mostly the different spatial footprints between ground and space-born observations that pose difficulties in evaluating and validating remote sensing rainfall estimates, especially for daily and sub-daily time periods.

Remote sensing estimates of precipitation can contain three types of errors: missed events (no-detection of events), false alarms (detection of rainfall not recorded on the ground), and errors in the rain-rate magnitude of correctly detected rainfall events. Different performance metrics can be used to quantify performance based on these types of errors. However, the errors from ground measurement networks also need to be considered and reference datasets need to be benchmarked [34]. It has been found that the true errors in satellite rainfall estimates are significantly lower when the errors in ground networks and the covariance between errors in the two types of observations are acknowledged [35]. If errors in the ground truth are not accounted for, analysis can inflate error estimates of SPPs [3,35]. While satellite products are indirect estimates of rainfall at the ground surface, heterogeneous precipitation over a gridded area may or may not occur over ground gauges and rainfall rates may significantly vary in space, representing a significant source of error for most ground-based networks, the magnitude of which depends on network density and rainfall system structure, storm cell size.

Thus, the best representation of the precipitation "ground truth" is one that assimilates both types of data: (1) direct point observations of rainfall at the ground surface; and (2) satellite estimates that provide the spatial distribution of rainfall. For that reason and for the continental scope of

5 of 22

this work, the GPCP-1dd was chosen as a reference dataset against which to compare the satellite precipitation products. As with every approach and choice of reference dataset, this one has some potential limitations that will be discussed below. In summary, any continental or global dataset integrating the best possible retrospective information from a range of in-situ measurements and satellite sensors will have some conflict of interest with the real-time satellite products evaluated. This conflict is minimized by evaluating near-real time estimates and not their improved versions that incorporate gauge data a posteriori. The alternative to using a global reference dataset is to constrain analysis to a location with a good rain-gauge and radar coverage, which is not the case of Africa. Without the use of a continuous reference dataset, one could not evaluate performance metrics in a continuous manner across space and time (different geographies and seasons) and connect this analysis with continental climate mechanisms and dynamics. The strength of our choice is that the GPCP-1dd is a well-established dataset that correctly captures the historical climatology of the African continent by integrating different sources of historical data for the period of analysis.

## 3. Materials and Methods

## 3.1. Study Area and Datasets

A domain covering the entirety of the African continent was selected, constituting the region between longitude 18°W to 52°E and latitudes 38°N to 35°S. The time period considered in this study is from January of 2001 to December of 2013, for which the three satellite products and the reference dataset are available. The three SPPs evaluated, originally at 0.25°, were aggregated to a resolution of 1° to enable a direct cell to cell comparison with the reference dataset (GPCP-1dd).

GPCP-1dd, our reference dataset, is the product of a careful combination of more than 6.500 rain gauge stations and high quality remotely sensed data, such as geostationary and low orbit infrared, passive microwave and sounding observations, which are not fully available for use in real time [36,37].

TMPA near real time product, 3B42-RT, uses passive MW rainfall estimates collected from the sensors TRMM Microwave Imager (TMI), Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), Special Sensor Microwave/Imager (SSM/I), and Advanced Microwave Sounding Unit B (AMSU-B), and processes the former three using the Goddard Profiling Algorithm. The highest quality estimate is the TRMM Combined Instrument Algorithm (TCI)—TMI plus the precipitation radar—used to calibrate the TMI estimates, which are then used to calibrate the SSM/I data. All the calibrations consist on climatological adjustments. With all the PMW data calibrated, a surface is populated with the "best" value available for each pixel. Then a high quality estimate is created by merging all the MW datasets. Merged IR estimates delivered by the CPC each 30 min are calibrated against the high quality MW estimate by matching both histograms average on coincident areas. TMPA 3B42-RT is produced with a latency of 3 h as a combination of both sources of data, using the best quality product available on each point [38].

CMORPH uses satellite IR brightness temperature from geostationary IR sensors, with a latency of 30 min to 3 h for any point on the globe, and a resolution of about  $0.04^{\circ}$ ; as well as PMW estimates from the sensors TMI, SSM/I, and AMSU-B. The spatial resolution of these MW estimates is  $4.6 \times 6.9$  km,  $13 \times 15$  km, and  $15 \times 15$  km respectively. The rainfall mapping is produced from the PMW satellites and advected or "morphed" using IR images. The cloud system advection vectors (CSAV) that determine the "morphing" are calculated by computing the spatial half-hour correlation of the time lagged IR images. A correction is applied to those CSAV's to account for a discrepancy between satellite and radar observed movement vectors in S-N and W-E directions. This correction was computed for the US but applied globally [39].

PERSIANN, developed at The University of Arizona and now operated by the CHRS, UC-Irvine, is based on an adaptive Artificial Neural Network (ANN) model that estimates precipitation using IR information from geostationary satellites to feed a Modified Counter Propagation ANN that has been trained with ground and MW satellite estimates to give an output of rainfall rate. Such ANN classifies

the input IR data plus other important information (topography, location, wind flow, etc.) using a pre trained self-organized feature map (SOFM), and then maps this and the input variables through an Input–Output Prediction Map to create an estimated rainfall rate. Other high quality rainfall estimates, such as rain gauges, precipitation radar and MW rainfall estimates are used to train and adapt the input-output prediction map parameters to improve further predictions. PERSIANN estimates have a latency of up to 1 day [40].

#### 3.2. Satellite Product Evaluation and Error Analysis

In order to evaluate SPPs in space and time, different temporal and spatial levels of analysis were used. Using daily data as the time unit of interest, all in UTC time, the analysis is done at seasonal and annual levels. Several spatial classifications have been used for this work:

All of Africa: Metrics were computed for the entire continent including: Probability of Detection (POD), False Alarm Rate (FAR), Equitable Threat Score (ETS) and Nash–Sutcliffe Efficiency Coefficient (NSE).

Latitudinal bands: Given that the ITCZ is the single main rain-producing mechanism in Africa, and because it oscillates around the equator, there is a clear trend between latitude and rainfall. Consequently, latitudinal bands were defined based on zones that have similar rainfall patterns and vegetation cover based on USGS Africa Land Cover Characteristic Data Base (Version 2.0). The domains of each latitudinal band are defined in Table 1 and shown in Figure 1. In each latitudinal domain, twelve different statistics were calculated for each SPP with respect to the reference dataset GPCP-1dd. This was done for four seasons corresponding to December, January, and February (DJF); March, April, and May (MAM); June, July, and August (JJA); and September, October, and November (SON). In addition, the POD, FAR and ETS were also evaluated for each product in each band and season.

**Climatic Classification**: The same as above was done for three general climates of the Koppen–Geiger climatic classifications for Africa: Arid, Warm and Tropical [41], as shown in Figure 1 (middle).

**Topography**: As described in the previous section, performance of SPPs in mountainous areas can be low when detecting orographic rainfall, and a distinction is made here between "flatlands" and "mountainous regions" to evaluate differences in performance in the two types of areas.



**Figure 1.** The African Continent with the three types of spatial analysis: (**a**) latitudinal bands (A to F, from North to South); (**b**) Koppen–Geiger climatic zones; and (**c**) topography.

Table 2 shows the statistics computed for the analysis over each spatial domain and seasonal aggregation. The POD, FAR and ETS reflect the ability of an SPP to correctly detect rainfall when it occurs. In addition, we can highlight three main error components: bias (errors in the mean); differences in variability (reflecting errors in representation of the observed variability); and correlation

(measuring the strength or closeness of the linear relationship between SPP estimates and the reference data). These error components can be mathematically expressed in a number of metrics, which vary across the literature. It is important to point out that the focus of the analysis is on the land areas and a mask has been applied to remove all the cells containing coastal waters and oceans, for the computation of aggregate statistics in latitudinal bands, climate zones, and topographic areas.

Band	Boundaries	Description				
A—Sahel	Sahara Desert. Sparse vegetation to bare areas					
B—Semi-Arid North	Transition between Sahara Desert and Savanna. Desert and grasslands					
C—North Savanna	Lat: 15°N–5°N Lon: 20°W–55°E	North Savanna. Croplands, grassland and open deciduous forest and scrubland				
D—Tropical zone	D—Tropical zone Lat: 5°N–5°S Lon: 20°W–55°E Evergreen and semi-deciduous fores					
E—South Savanna	Lat: 5°S–15°S Lon: 20°W–40°E	Open and close deciduous forests, scrubland and grasslands				
F—Semi-Arid South	Lat: 15°S–40°S Lon: 20°W–40°E	Grasslands, croplands				

 Table 1. Latitudinal bands domain boundaries and land cover observations.

Table 2. Statistics computed for each of the spatial domains and season.

Statistic	Description	Units
POD	Probability of Detection: Hits/(Hits + Misses)	-
FAR	False Alarm Rate: $FA/(H + FA)$	-
ETS	Equitable Threat Score: $(H-Hrand)/(H + M + FA - Hrand)$ Where Hrand: $[(H + M) \times (H + FA)]/T$	-
Mean_sat.	Temporal mean of SPP over the spatial extent	mm/day
Mean_grnd	Temporal mean of the reference precipitation over the spatial extent	mm/day
pBias	Percent bias of the SPP with respect to the reference precipitation	-
Correl	Correlation coefficient between SPP and reference precipitation	-
std_mon	Standard deviation between both products	mm/day
RMS_mean	Part of the RMS corresponding to error in the mean	mm
RMS_var	Part of the RMS corresponding to error in the variance	mm
RMS_corr	Part of the RMS corresponding to error in the correlation	mm
RMS	Root mean squared error	mm
NSE	Nash–Sutcliffe Efficiency	-
MaxsAT	Maximum SPP value	mm/day
MaxRP	Maximum reference value	mm/day

In order to gain understanding on the sources and nature of errors in SPPs, we have used a decomposition of the Mean Square Error (MSE) as the sum of three error terms: error in the mean, variance and correlation, as shown in Equation (1) [42].

$$MSE = 2\sigma_s \sigma_o (1 - r) + (\sigma_s - \sigma_o)^2 + (\mu_s - \mu_o)^2$$
(1)

where  $\sigma_S$ ,  $\sigma_o$  are the standard deviations of the satellite and the reference dataset, respectively; *r* is the correlation coefficient between satellite and reference data; and  $\mu_s$ ,  $\mu_o$  are the mean of the satellite and ground truth data, respectively.

#### 3.3. Bias Correction and Error Reduction Analysis

A quantile mapping approach was used for bias correction in this work, where the distribution functions of the satellite estimates in each latitudinal band domain were adjusted to the distribution of the observations in the following way:

$$Pp_{corr} = F_{ref}^{-1} \left( F_{SPP} \left( Pp_{SPP} \right) \right) \tag{2}$$

where  $Pp_{corr}$  is the bias corrected precipitation,  $F_{ref}^{-1}$  is the inverse of the reference dataset cumulative density function (CDF),  $F_{SPP}$  is the CDF of the SPP estimates, and  $Pp_{SPP}$  is the rainfall estimates from the SPP. After bias-correction, the error analysis is repeated for the new estimates, and a Mann–Kendall test is used to determine significant trends between the magnitude of the errors and that of the rainfall intensity, and the presence of heteroscedasticity.

# 3.4. Product Averaging

Once SPP estimates have been bias corrected and the mean and standard deviation error terms are expected to be significantly reduced, the averaging of product estimates can, under certain circumstances, reduce the correlation error term. Winter and Nychka (2010) define such circumstances as occurring when estimates have a negative correspondence of errors and can partially cancel each other [43]. This condition can be expressed as follows:

$$\sum_{m \neq m'} (S_m \times S_{m'} \times -S_{min}^2) + \sum_m (S_m^2 - S_{min}^2) < 0$$
(3)

$$S_m \times S'_m \times \cos(\theta) = \frac{1}{N} \sum_{t=1}^N \epsilon_{m,t} \times \epsilon_{m',t}$$
(4)

where the first sum is the measurement of the correspondence between errors,  $S_m^2$  is the MSE of the model *m* (*m*' denotes a different model from m),  $S_{min}^2$  is the MSE of the best model or minimum MSE,  $\epsilon_{m,t}$  is the error of the model m at the time *t*, and  $\theta$  is the angle between error vectors of *m* and *m*'. By checking this condition for the SPP estimates considered here, we can predict if the average of estimates will perform better than the best individual estimate.

#### 4. Results and Discussion

#### 4.1. An Africa-Wide Analysis

The average annual precipitation for the period 2001 to 2013, estimated by the SPPs—TMPA 3B42-RT, CMORPH and PERSIANN—is shown in Figure 2, as well as for the GPCP-1dd reference dataset. It can be observed that the greatest difference between SPP estimates is in the Central Africa region. In the absence of ground gauges to help calibrate algorithms in that area, the difference is due to the fact that while CMORPH and PERSIANN only assimilate information from IR and MW, TMPA also incorporates information from TRMM's orbital radar, thus yielding different estimates. As seen in Figure 3, the number of rain-gauges available for the reference dataset has a very poor coverage in the Central Africa region. The NSE Coefficients between each one of the SPPs and the GPCP-1dd reference dataset (Figure 4) reflect a potential conflict of interest between the latter and the TMPA product. TMPA has higher NSE values over the entire continent but especially in the Central Africa region where the other products have low or negative NSE values (over DR Congo, Central African Republic, South Sudan and Angola). This difference could be due to the fact that the reference dataset does not have ground observations in Central Africa and thus relies more on the TRMM radar, one of the main data sources of the TMPA product. Thus, TMPA does significantly better in that region while CMORPH and PERSIANN (which do not use TRMM data) do worse in this comparison. This may illustrate a conflict of interest between the TMPA 3B42-RT and the reference dataset, especially over that region, where the reference dataset heavily relies on TRMM, one of the main inputs to TMPA. However, CMORPH and PERSIANN also do worse over neighboring regions with rain gauges, such as Republic of Congo, Gabon, Western part of Central African Republic, Southern Chad, and further to the West over Ghana, Togo and Benin. On the other hand, CMORPH and PERSIANN do significantly better than TMPA over the Horn of Africa region, where no rain gauges are used in the GPCP product. Thus, the potential conflict of interest between TMPA and the reference dataset due to their use of TRMM radar data does not necessarily influence the performance of TMPA with respect to GPCP.



Over the Horn of Africa region, results show that IR and PMW data used in PERSIANN and CMORPH may be providing the core of the information for the GPCP estimates over the region.

**Figure 2.** Average annual rainfall (in millimeters) for the period of study as seen by the reference dataset GPCP-1dd (Global Precipitation Climatology Project one Degree Day), and the three products TRMM Multi-satellite Precipitation Analysis (TMPA), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN), and CPC Morphing Technique (CMORPH).



**Figure 3.** Number of ground gauges per cell over Africa, considered in the GPCC. As can be seen, the region over central Africa (DR Congo, Central African Republic, South Sudan and Angola) stands out as a big void for its lack of reporting rain-gauges (source: GPCC).



**Figure 4.** Nash–Sutcliffe Efficiency (NSE) of the three Satellite Precipitation products with respect to the reference database GPCP-1dd before bias correction (top row), after bias correction (middle row) and the relative improvement (bottom row).

NSE values deteriorate quickly for the three products beyond the area of influence of the ITCZ, and are worse between 15°N and 30°N, over the Saharan Desert. This significant decrease in performance, mirrored by POD and FAR (see Figure 5), is mainly due to the scarce rainfall and its high spatial variability at all-time scales. Ali et al. (2005b) shows how global products tend to highly underestimate the frequency of small rain events, linked to the high spatial intermittency at the rain event scale, which may better be captured by point measurements [3]. While TMPA does better in the Northern part of the Sahara between 20°N and 30°N, CMORPH and PERSIANN do slightly better in the region between 15°N and 20°N. NSE values remain positive and relatively high—especially for TMPA—over all of Southern Africa, down to 32°S. The region where CMORPH and PERSIANN have higher NSE values is along the Eastern Coastal countries of Africa, from Ethiopia to Mozambique.

Bias correction significantly improves NSE values for CMORPH and PERSIANN across the continent south of latitude 20°N, Efficiency values for CMORPH and PERSIANN over some regions in the arid areas of the Sahara and the Horn of Africa very slightly decrease after bias correction, likely due to an artifact of bias correction in these arid regions with erratic rainfall events. As shown in the relative improvement of Figure 4, the bias correction increases NSE over most of the continent below 20°N for CMORPH and PERSIANN but the highest improvements for TMPA are limited to the southern fringes of the Sahel with a slight increase in NSE across the continent, and mostly between 20°N and the equator. It is worth noting that PERSIANN has the highest NSE values over the largest extension of the African continent after bias correction, due to its better efficiency in the fringe between 15 and 20°N. Notably, the bias correction does little to improve values over most of the Saharan Desert and North Africa, which is an indication that most of the error over the region is a correlation error, related to the difficulty of correctly detecting rain.



**Figure 5.** Probability of Detection (POD) and False Alarm Rate (FAR) rates of the three Satellite Precipitation products with respect to the reference database GPCP-1dd over Africa.

The Probability of Detection is high for the three products except in the Saharan Desert, the Horn of Africa, and Southern Africa (Namibia, Botswana, and South Africa) region (Figure 5). While TMPA shows slightly higher POD values over Central Africa and Western sub-Saharan Africa PERSIANN shows a clear superiority detecting rain events over the Horn of Africa and East Africa regions. FAR values are lowest over the ITCZ area of influence. Over the Horn of Africa FAR values increase for all products, with PERSIANN showing the biggest spatial extent of higher False Alarm Rates over that region. Both POD and FAR values are best over the continent below 15°N, but while POD deteriorates below 15°S, especially for TMPA and CMORPH, FAR values remain low (good) beyond the area of influence of the ITCZ until beyond 30°S. The decrease in event detection with increasing aridity is in agreement with previous findings [13,14].

This change in performance beyond the southern boundaries of ITCZ influence is not reflected in NSE or FAR values. The overall better performance of SPPs beyond the ITCZ influence to the South, as compared to the North (Saharan Desert), is due to the influence of the Indian Ocean and the East African Monsoon, which result in the dominance of convective rainfall over the region, which SPPs estimate best.

The POD and the FAR of a rainfall product are the result of several factors, including the calibrated relationships between sensor reflectivity and estimated rain-rate, and the sampling error depending on the timing of satellite overpasses. They are related metrics to a certain extent, and their relationship depends on the algorithms and procedures used in the product development.

#### 4.2. Latitudinal Band Analysis

The dependency of satellite precipitation errors on the dynamics of rain producing mechanisms—specifically the ITCZ over the African continent—is illustrated at the seasonal level in Table 3, which shows the relative errors of each SPP for each season and for each latitudinal band as defined in Figure 1. The cell color indicates higher (red) and lower (green) errors. The lower errors are always in the bands where the ITCZ is located during a specific season. Smaller relative errors where rainfall mechanisms are convective, frequent and regional is in agreement with findings from

Artan et al. (2007), who attributed the good results in simulations of Mekong sub-basins during the monsoons to those characteristics [44].

**Table 3.** Seasonal Relative Errors across latitudinal bands for the three products. It can be seen how the smallest relative errors follow the oscillation of the Intertropical Convergence Zone (ITCZ), indicated by bold borders in cells.

Band -		TM	PA			СМО	RPH		PERSIANN				
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	
А	1.7	1.2	1.2	1.2	2.9	1.8	1.3	1.6	2.4	1.7	1.4	1.3	
В	2.2	1.4	0.6	1.0	3.0	1.7	0.9	1.3	2.8	1.9	0.8	1.2	
С	0.7	0.2	0.2	0.2	1.9	0.8	0.7	0.8	1.1	0.7	0.5	0.6	
D	0.4	0.2	0.3	0.2	0.3	0.7	0.8	0.7	0.7	0.6	0.6	0.6	
Е	0.2	0.3	1.0	0.4	0.3	0.9	2.1	0.8	0.5	0.6	2.0	0.8	
F	0.3	0.5	1.2	0.5	0.8	1.3	3.8	1.0	0.4	0.8	11.6	1.0	

- Band A: In general, none of the products show a good performance in this area. The ETS (Equitable Threat Score) shows no skill in the detection of rain, poor POD and FAR values. These results are mostly due to the lack of regional convective rainfall activity, and the difficulty of SPPs in detecting frontal rainfall processes. They are in agreement with previous results reported in the literature for such latitudes [3].
- Band B: This fringe is a transition between the Sahel and the zone influenced by the ITCZ. Rainfall occurs mostly between June and August, coinciding with the movement of the ITCZ, but there are rainfall events up to November. All the products show a strong positive bias during the rainy season that is successfully removed with the bias correction. An interesting feature of the PERSIANN data is that is seems to be more accurate than TMPA and CMORPH, but it also shows more frequent non detected precipitation events than the other products. It has better skill on estimating the rain-rate if correctly detected, but a higher failure rate in detecting it. The detection skill (ETS) of all SPPs is low, with better performance in JJA, when most of the rain occurs.
- Band C: This latitudinal band is where the SPPs show the best overall skill in observing rainfall. Most of the rainfall occurs between June and August, but with significant rainfall in the seasons before and after. CMORPH and PERSIANN overestimate annual rainfall by nearly 50%. The bias correction has a significant impact, improving the NSE from the vicinity of -1 to 0.78–0.95. TMPA has the least bias of the three SPPs, and its NSE ranges between 0.7 and 0.9 that is increased to 0.88–0.95 after the bias correction. It is also in this band where the errors seem less dependent of the magnitude of rainfall, but heteroscedastic nevertheless, as shown by the Mann–Kendall tests. The POD, FAR and ETS are good and relatively similar between products during the rainy seasons, but the detection on the dry season is still poor.
- Band D: This tropical band is characterized by two rainy seasons due to the double transit of
  the ITCZ over the equator during the spring and fall of the Northern hemisphere. The error
  characteristics are very similar throughout the year: TMPA underestimates less than 10% in
  all seasons, while CMORPH and PERSIANN overestimate over 30% of the actual rainfall.
  After correction, the bias is reduced to less than 5%, except for PERSIANN, where the missed
  rainfall events seem to affect the overall good performance of the hits. This latitudinal band
  shows the best performances in rain detection for all seasons, confirming past findings that SPPs
  work best detecting tropical convective storms. The higher performance of SPPs in tropical and
  equatorial regions with strong convective activity has been widely reported [7,10,14–16].
- Band E: Mirroring band C, most of the annual rainfall occurs between December and February, with significant precipitation on the seasons before and after. Biases are similar to those of bands C and D, with underestimation of TMPA and overestimation of PERSIANN and CMORPH, but with highest bias between September and November. The errors after the correction seem to be greater

than those on band C. Again, PERSIANN has the least dispersion between SPP estimates and observations, but many missed rainfall events. The rain detection skill during the dry season is notably poorer than during other seasons, and the best performance is between December and February, when the ITCZ is over the area.

• Band F: In this zone, the bias is positive on the rainy months and negative on the dry season. Precipitation occurs mostly between December and February that is only partially removed by the bias correction.

The effects of bias-correction are also shown in Figure 6 for seasonal precipitation of PERSIANN and latitudinal band D. Most notably, the mean and standard deviation error terms have been almost eliminated, but have a less significant effect on the correlation error term. The figure is representative of other latitudinal bands.



**Figure 6.** Scatterplots of daily precipitation of PERSIANN and GPCP in latitudinal band C before (**top**) and after (**bottom**) bias correction. The effect of bias correction is shown in the performance statistics and in the decomposition of the mean square error components. (These figures with their corresponding statistics tables are available for all three products, latitudinal bands, climatic and topographic zones, before and after bias correction, as online Supplementary Materials in Figures A1 to F6).

## 4.3. Climatic and Topographic Classification Analysis

In regards to the Koppen Geiger Climatic zone classification, SPPs work the best in the tropical zones due to the convective nature of rainfall. Raw CMORPH and PERSIANN clearly overestimate rainfall by 30% and 40%, respectively, and TMPA slightly underestimates it by about 3% on average (not shown here, see climate classification scatterplots and tables in the Supplementary Materials online). The bias correction improves the estimates, but not nearly as much as with the latitude band groupings. The rather constant values of POD, FAR and ETS through the year indicate poor isolation of specific rain regimes. In general, the best skill is for tropical zones, followed by arid zones. For warm zones the performance is poor because of the very high FAR, even if TMPA and CMORPH show small bias while PERSIANN has a strong negative bias between March and August, and a positive bias the

rest of the year. On arid zones the products overestimate 5% to 50% of the rainfall, a bias that is partially removed by the correction, yielding a mixture of overestimations up to 10% and underestimations down to -30%. The NSE value does not improve much for TMPA, but it does for CMORPH and PERSIANN between September and February, from negative values to low positive values.

While the NSE of plains is consistently better than that of mountainous areas, the lack of other significant differences between flat and mountainous areas suggest that the topographic classification used here, grouping distant mountainous areas across the continent, at different latitudes and with different rainfall regimes and climatology (and at the scale of one degree), is likely smoothing out other differences that could be better characterized locally through limited-area case studies.

# 4.4. Characterizing Biases and the Effect of Bias Correction

The effects of bias correction on the MSE error terms (mean, variance, correlation) and the NSE values can be seen for each spatial area (latitudinal band, climate and topographic) for a specific season (Tables 4–6). The statistics are calculated for the daily spatial average over each area. Results show that bias correction is generally very successful at removing the mean error term of MSE as well as a significant portion of the variance error term. However, since the correlation term depends on the lack of co-fluctuation between estimates and observations, bias correction does little to reduce it. In the majority of the cases, for all seasons and products, the Nash–Sutcliffe Efficiency values improve with bias correction, especially for CMORPH and PERSIANN. The climatic warm areas are an exception to this improvement, where bias correction does not improve NSE values and in almost 50% of the cases, it worsens it. The likely reason for this is the spotty and uneven distribution of warm areas on the northern and southern hemispheres. By grouping these areas together, the bias correction mixes different rainfall regimes, thus losing its efficiency. For better results, the bias correction should be performed separately in each warm area.

Our approach to delimit spatial areas of analysis and seasonal evaluations is based on the assumption that such grouping of data allows us to isolate rainfall regimes and climatic controls for which SPPs may exhibit specific error characterizations. In cases where that holds true, bias correction can successfully remove or reduce significant parts of the error. On the other hand, if the grouping is unsuccessful, and several rainfall regimes are mixed together, as may be the case for climatic warm areas, bias correction will be challenged by the presence of data samples with errors of different nature. From this analysis, it becomes apparent that latitudinal bands are the large areas of analysis that better capture the seasonality of rainfall regimes over Africa, as they align well with the dynamics of the ITCZ.

TMPA JJA		Latitude Bands							Climate Class			Topo. Class	
		Α	В	С	D	Ε	F	Arid	Warm	Tropical	Plains	Mtn.	
Avg. GPCP (mm/day)		0.10	1.14	5.19	2.35	0.25	0.25	1.06	1.17	4.37	1.79	2.94	
Mean Error Term of MSE	Raw BC	0.00 0.00	0.28 0.00	0.09 0.00	0.06 0.00	0.00 0.00	0.00 0.00	0.02 0.01	0.00 0.00	0.00 0.01	0.01 0.01	0.01 0.05	
Variance Error Term of MSE	Raw BC	0.00 0.00	0.12 0.01	0.00 0.02	0.04 0.05	0.00 0.00	0.03 0.02	0.00 0.00	0.02 0.01	0.00 0.02	0.00 0.00	0.0 0.00	
Correlation Error Term of MSE	Raw BC	0.02 0.01	0.32 0.23	0.93 0.87	0.47 0.43	0.06 0.06	0.08 0.09	0.09 0.08	0.23 0.24	0.94 0.86	0.13 0.12	0.62 0.66	
NSE value	Raw BC	0.39 0.54	0.38 0.80	0.74 0.78	0.72 0.77	0.45 0.46	0.45 0.46	0.71 0.76	$0.54 \\ 0.54$	0.71 0.73	0.71 0.73	0.59 0.62	

**Table 4.** Mean Square Error (MSE) Terms and Nash-Sutcliffe Efficiency (NSE) before (Raw) and after bias correction (BC) for TMPA estimates, June to August (JJA) season.

		Latitude Bands						Climate Class			Topo. Class	
CWORTHJJA	CWOKFHJJA		В	С	D	Ε	F	Arid	Warm	Tropical	Plains	Mtn.
Avg. GPCP (mm/day)		0.10	1.14	5.19	2.35	0.25	0.25	1.06	1.17	4.37	1.79	2.94
Mean Error	Raw	0.02	0.16	2.67	0.40	0.00	0.01	0.12	0.00	1.48	0.24	0.47
Term of MSE	BC	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.11	0.16	0.03	0.08
Variance Error	Raw	0.03	0.03	0.02	0.07	0.01	0.06	0.01	0.03	0.01	0.00	0.00
Term of MSE	BC	0.00	0.00	0.13	0.06	0.00	0.00	0.01	0.08	0.25	0.03	0.09
Correlation Error	Raw	0.02	0.33	1.55	0.87	0.10	0.09	0.14	0.37	1.28	0.20	0.90
Term of MSE	BC	0.01	0.28	1.17	0.60	0.08	0.17	0.10	0.30	0.86	0.14	0.65
NSE value	Raw	-1.10	0.54	-0.06	0.35	0.07	0.24	0.29	0.25	0.14	0.10	0.11
	BC	0.70	0.75	0.67	0.68	0.29	0.83	0.71	0.10	0.61	0.60	0.47

**Table 5.** Mean Square Error (MSE) Terms and Nash-Sutcliffe Efficiency (NSE) before (Raw) and after bias correction (BC) for CMORPH estimates, June to August (JJA) season.

**Table 6.** Mean Square Error (MSE) Terms and Nash-Sutcliffe Efficiency (NSE) before (Raw) and after bias correction (BC) for PERSIANN estimates, June to August (JJA) season.

PERSIANN JJA		Latitude Bands							Climate Class			Class
		Α	В	С	D	Ε	F	Arid	Warm	Tropical	Plains	Mtn.
Avg. GPCP (mm/day)		0.10	1.14	5.19	2.35	0.25	0.25	1.06	1.17	4.37	1.75	2.94
Mean Error Term of MSE	Raw BC	0.01 0.00	0.89 0.00	5.31 0.02	0.76 0.05	0.00 0.01	0.05 0.02	0.36 0.00	0.21 0.32	1.92 0.23	0.57 0.00	0.10 0.23
Variance Error Term of MSE	Raw BC	0.02 0.00	0.78 0.00	0.98 0.00	0.46 0.01	0.04 0.00	0.11 0.02	0.16 0.01	0.03 0.08	0.36 0.03	0.11 0.00	0.10 0.01
Correlation Error Term of MSE	Raw BC	0.02 0.01	0.37 0.21	2.91 1.95	1.31 0.84	0.12 0.07	0.05 0.13	0.18 0.12	0.33 0.26	2.42 1.62	0.38 0.26	1.21 0.89
NSE value	Raw BC	-0.43 0.60	-0.77 0.82	-1.73 0.51	-0.23 0.57	$-0.33 \\ 0.32$	0.00 0.21	-0.84 0.66	$-0.04 \\ -0.23$	$-0.45 \\ 0.42$	$-1.19 \\ 0.46$	0.09 0.27

The average bias of daily rainfall estimates is shown for each of the three SPPs in Figure 7 as a function of rainfall intensity, for each latitudinal band. The bias for a given rainfall intensity has been computed by averaging estimates from "hit" events in all the cells within the latitudinal band, before bias correction. The red histogram in the figures represents the relative frequency of errors within the band. The dependency of bias on rainfall intensity is different for the three products. Overall, the bias of CMORPH and PERSIANN is less dependent on rainfall intensity, with a slightly increasing bias with higher rain-rates. In bands C to E where it rains the most, CMORPH bias factors can range between 0.8 (+25% bias) and 0.6 (+66%), with an overall increasing bias. PERSIANN exhibits similar trends, but with slightly different shapes such as flatter line for most intensities, resulting in higher biases for lower intensities, and a slight increase in bias with increasing intensities (i.e., a decrease in bias factors). Because the Saharan Desert and North Africa is where all products perform the worst and hit events are fewer, the relationship is noisier than in the other bands but similarly shaped in the two products. TMPA estimates have a clear dependency on rainfall intensity. Bias factors have a clear peak between intensities of 5 and 10 mm/day, which in bands C, D, and E is around 1.2 (-16%). As intensities increase, the bias increases up to values of 33% and 66% (with bias factors decreasing to 0.75 and 0.6 respectively). In general, all SPPs have relatively similar biases for mid and high rain intensities, with TMPA showing lower biases at low rain intensities. All products overestimate high rain intensities, as found by others [11].



**Figure 7.** Bias correction factors for the three products (solid lines), and their relative frequency of daily biases (dotted lines) as a function of rain intensity within each latitudinal band.

The effect of the bias correction at specific locations was also analyzed looking at time series of rainfall errors before and after the correction, in cells of each latitudinal band. Figures 8–10 display the time series of raw and bias corrected estimates in a representative cell within bands C (upper plot), D (middle) and E (lower), for each SPP. The errors show CMORPH and PERSIANN's general tendency towards overestimation of rainfall, in the three bands. Coherent with TMPA's varying biases depending on intensity, Figure 10 shows both over and underestimation errors. The positive bias is more persistent in PERSIANN and the error is on the side of overestimation for the great majority of days. Because the variance is smaller, the final residual series follow closely the observed data and show slightly lesser errors than TMPA and CMORPH, for these three specific cells. After bias correction, the error distributions are symmetrical with zero mean.



Figure 8. Errors of raw and bias corrected CMORPH data over representative cells in bands C, D and E.



**Figure 9.** Errors of raw and bias corrected PERSIANN data over representative cells in bands C, D and E.



Figure 10. Errors of raw and bias corrected TMPA data over representative cells in bands C, D and E.

Using the bias corrected estimates, the daily errors over each pixel were grouped for each season, SPP and analysis area (latitude bands, climatic zone, and topography type). The error distribution was symmetric with zero mean, and because of that, the absolute value of the error was taken. The Mann–Kendall tests for each product, latitudinal band and season, found positive trends between the magnitude of the errors after bias correction and the magnitude of the rainfall in all products, bands and seasons, except in one case (band C, JJA season, PERSIANN). The existence of the trend indicates that the magnitude of the errors vary with the magnitude of the rainfall, implying a non-constant variance and heteroscedasticity.

The absolute values of the rain-rate errors (considering "hits" only: when rain occurs and is detected by the SPP) after bias correction can be approximated by a Gamma distribution. Table 7 shows the median rain-rate errors considering the "hit" cases only. Minimum values are marked in bold. While TMPA and CMORPH have shown better correlation coefficients, NSE, and MSE values than PERSIANN, it can be seen that the median error of PERSIANN is usually smaller than for the other two products. This can be explained by the fact that only "hit" events are used to compute the histogram of errors: days where rain occurred and the SPPs correctly detected it. This leaves

out missed events, false alarms, and correct negatives. Thus, it can be stated that while it has higher detection errors as seen in the x-axis of Figure 6, PERSIANN is shown here to be more accurate in the rain-rate estimation once it correctly detects rainfall.

DJF	Median	Rain-Rate Erro	r (mm/Day)	JJA	A Median Rain-Rate Error				
Band	TMPA	CMORPH	PERSIANN	Band	TMPA	CMORPH	PERSIANN		
А	1.5	1.2	1.2	А	1.0	0.8	0.8		
В	0.3	0.2	0.2	В	2.9	2.6	1.7		
С	0.9	0.7	0.9	С	2.9	3.0	2.4		
D	2.4	2.5	1.8	D	1.9	1.9	1.4		
Е	2.6	2.9	1.7	E	1.0	0.8	0.8		
F	2.6	2.7	1.7	F	1.4	1.0	1.7		
MAM	Me	dian Rain-Rate	e Error	SON	Median Rain-Rate Error				
Band	TMPA	CMORPH	PERSIANN	Band	TMPA	CMORPH	PERSIANN		
А	0.9	0.7	0.7	А	1.7	1.3	1.0		
В	0.8	0.5	0.3	В	2.0	1.6	0.8		
С	2.1	2.1	1.6	С	2.3	2.3	1.8		
D	2.3	2.5	1.8	D	2.4	2.6	2.0		
Е	2.1	2.3	1.8	Е	1.8	2.0	1.5		
F	1.8	1.9	1.7	F	1.7	1.7	1.4		

**Table 7.** Median rain-rate error after bias correction for each SPP, season and latitudinal band, considering hits only.

# 4.5. On Product Averaging

In order to determine if the average of several SPPs has more skill than the most skilled product, the condition from Winter and Nichka described in the Methods Section was computed for each pixel in the domain. Figure 11 shows the few locations where averaging was found to perform better than the best individual product. The colors indicate the number of seasons in which this condition holds true. Product averaging shows promise on the Sahel region, especially over Egypt. In general, the satellite precipitation products used here are too similar to present the negative correspondence property that makes averages better than the individual products.



**Figure 11.** Locations where the arithmetic average of SPPs is a better estimate than any individual product. The color code indicates the number of seasons where this holds true, with blue for 1 season, green for 2 seasons and orange for 3 seasons.

In a pragmatic and original approach to benefit hydrologic applications in un-gauged regions where SPPs are difficult to evaluate due to lack of reference datasets, Gebregiorgis and Hossain (2014) develop a regression error model between runoff error variance and satellite precipitation estimates that enables the weighting of satellite products to produce a merged rainfall product [45]. The new estimate should yield improved runoff error variance and better performance in hydrologic applications. In regions with ground observations the parameters of the regression are calibrated minimizing the runoff error variance using the output of model simulations driven by ground rainfall as runoff observations. The parameters can then be extrapolated to un-gauged regions—based on climate and topography similarities—to calculate the runoff error variance, which in turn is then used in each cell to weight and merge different estimates. Of all products, the merged estimate showed the highest POD in all regions, a high FAR (but not highest), comparable threat score, and consistently lower RMSE, thus the approach effectively merges some products' higher ability to detect rain of the less biased rain rate estimates of other products.

#### 5. Conclusions

A daily error and bias correction analysis of three quasi-global satellite precipitation products across the African continent is presented here at a seasonal level for latitudinal bands, climatic regions and topography. Because of the main influence of the ITCZ, a rain-producing mechanism with an East–West alignment oscillating around the equator, and the East African Monsoon, latitudinal bands are found to capture well the changing nature of SPP errors across latitude and seasons. As widely established, satellite estimates do best in tropical and sub-tropical regions with convective precipitation regimes, both in detecting rain, as well as in terms of relative errors. Highlighting the strong dependency of rainfall errors on precipitation dynamics, the areas with smallest relative errors follow the position of the ITCZ across the seasons. The lowest performance for all products occurs over the Saharan Desert between 15°N and 30°N due to the difficulties in correctly detecting rain, as shown by low POD and high FAR values.

The bias correction analysis provides insights as to what error components can be reduced over what regions. The bias correction approach of adjusting rainfall distributions in each latitudinal band is very successful in removing the mean and variance error terms, but as expected does not correct the correlation error. While it significantly increases efficiency values over the continent below 15°N latitude, especially for CMORPH and PERSIANN, it has no effect on errors over the Saharan Desert to the North, as these arise from difficulties in detecting rainfall. The bias correction in areas delimited through the climatic and topographic classification did not yield good results due to the fact that such delimitations did not successfully isolate rainfall regimes. Individual case studies have better characterize errors in complex terrain.

Bias of TMPA estimates in each latitudinal band showed a higher dependency on rainfall intensity, going from negative values for lower intensities (<20 mm/day) to positive bias for higher intensities (>20 mm/day) in the tropical and subtropical latitudinal bands. CMORPH and PERSIANN were found to consistently overestimate rainfall but showed little dependency on rainfall rate, suggesting that a simple average bias correction may yield similar results as a pdf matching approach.

While TMPA and CMORPH have better correlations, NSE and MSE values, PERSIANN has generally the smallest rain-rate errors when only hits are considered, suggesting it is better at estimating the magnitude of the rainfall, but worse at correctly detecting precipitation events. While product averaging shows promise in some areas of the Saharan Region and Egypt, the products generally do not have the condition of a negative error correspondence for product averaging to work across the continent.

The reference dataset used in this study assimilates historical TRMM data, thus having a conflict of interest with TMPA, which is perhaps the reason for the higher NSE of TMPA raw estimates. However, interesting results such as the high bias dependency of TMPA on rainfall rates, and the lower median error of PERSIANN's hits, as well as the higher NSE of PERSIANN after bias correction, suggest this

conflict is not a major one. Since this study was completed, higher resolution reference datasets such as CHIRPS (5 km) have become available for use, with the caveat of that CHIRPS was developed for drought monitoring thus adjusted to match pentads and dekads, and not daily rainfall.

The error characterization of three SPPs over the African Continent, in relationship to latitude, topography, climate and season, allows for a better understanding of the relationships between the main climatic mechanisms controlling rainfall in Africa, such as the ITCZ; as well as the effects of bias correction approaches. Real-time SPPs constitute the best available tool to monitor spatial and temporal rainfall dynamics in poorly gauged basins and across Africa. Bias correction of real-time estimates is an essential step to improve their value and usability, and should be included in any application or research effort. A bias correction procedure that adjusts probability distributions of SPP to those of a reference dataset will take into account the biases of different sign and magnitude present for different rainfall intensities. The differences between SPPs are greatly reduced after bias correction, making them more consistent and usable for hydrologic monitoring applications at their current resolutions. Streamflow prediction for water allocation and reservoir operations in medium to large poorly-gauged basins, can greatly benefit from bias corrected real-time SPPs. While the existence of multiple SPPs is beneficial to the scientific community, their combined use can be advantageous for hydrologic applications such as the above. We hope the results presented here will continue to inform hydrologic applications of SPPs in Africa.

**Supplementary Materials:** The following are available online at www.mdpi.com/2072-4292/8/10/836/s1. Scatterplots and the values of error statistics shown in Table 2 of the manuscript for each satellite precipitation product, each season and each latitudinal band, climatic and topographic zones. Figure A-1–A-18: Raw Estimates Error Analysis for Latitudinal Bands. Figure B-1–B-9: Raw Estimates Error Analysis for Koppen Geiger Climatic Zones, Figure C-1–C-6: Raw Estimates Error Analysis for Topographic Classification Zones, Figure D-1–D-18: Bias Corrected Estimates Error Analysis for Latitudinal Bands, Figure E-1–E-9: Bias Corrected Estimates Error Analysis for Koppen Geiger Climatic Classification Zones, Figure F-1–F-6: Bias Corrected Estimates Error Analysis for Topographic Classification Zones.

Acknowledgments: The work of the second author was supported by Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) and GeoAguas Consultores (Chile). Support for the other authors was provided by the NASA SERVIR Program through award 11-SERVIR11-0058 and the International Center for Integrated Water Resources Management (ICIWaRM) a Category II UNESCO Center. All contributions are gratefully acknowledged.

**Author Contributions:** A.S.C., M.M. and J.V. conceived and designed the analysis; M.M. performed the calculations; M.M., A.S.C., J.V., and M.D. analyzed the data; M.D. provided data support; A.S.C. and M.M. wrote the paper; and J.V. and M.D. provided edits.

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

#### References

- 1. Shiklomanov, A.I.; Lammers, R.B.; Vorosmarty, C.J. Widespread decline in hydrological monitoring threatens panarctic research. *EOS Trans.* **2002**, *83*, 16–17. [CrossRef]
- 2. Stokstad, E. Scarcity of rain, stream gages threatens forecasts. *Science* 1999, 285, 1199–2000. [CrossRef]
- 3. Ali, A.; Amani, A.; Diedhiou, A.; Lebel, T. Rainfall estimation in the Sahel. Part II: Evaluation of rain gauge networks in the CILSS countries and objective intercomparison of rainfall products. *J. Appl. Meteorol.* 2005, 44, 1707–1722. [CrossRef]
- 4. Hughes, D.A. Comparison of satellite rainfall data with observations from gauging station networks. *J. Hydrol.* **2006**, 327, 399–410. [CrossRef]
- 5. Roca, R.; Chambon, P.; Jobard, I.; Kirstetter, P.E.; Gosset, M.; Bergès, J.C. Comparing satellite and surface rainfall products over West Africa at meteorologically relevant scales during the AMMA campaign using error estimates. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 715–731. [CrossRef]
- 6. Jobard, I.; Chopin, F.; Berges, J.C.; Roca, R. An intercomparison of 10-day satellite precipitation products during West African monsoon. *Int. J. Remote Sens.* **2011**, *32*, 2353–2376. [CrossRef]
- 7. Habib, E.; ElSaadani, M.; Haile, A.T. Climatology-focused evaluation of CMORPH and TMPA satellite rainfall products over the Nile Basin. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 2105–2121. [CrossRef]

- Gosset, M.; Viarre, J.; Quantin, G.; Alcoba, M. Evaluation of several rainfall products used for hydrological applications overWest Africa using two high-resolution gauge networks. *Q. J. R. Meteorol. Soc.* 2013, 139, 923–940. [CrossRef]
- 9. Guilloteau, C.; Gosset, M.; Vignolles, C.; Alcoba, M.; Tourre, Y.; Lacaux, J.P. Impacts of satellite-based rainfall products on predicting spatial patterns of rift valley fever vectors. *J. Hydrometeorol.* **2014**, *15*, 1624–1635. [CrossRef]
- 10. Diem, J.E.; Hartter, J.; Ryan, S.J.; Palace, M.W. Validation of satellite rainfall products for western Uganda. *J. Hydrometeorol.* **2014**, *15*, 2030–2038. [CrossRef]
- 11. 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]
- 12. Dinku, T.; Chidzambwa, S.; Ceccato, P.; Connor, S.J.; Ropelewski, C.F. Validation of high resolution satellite rainfall products over complex terrain. *Int. J. Remote Sens.* **2008**, *29*, 4097–4110. [CrossRef]
- Dinku, T.; Connor, S.J.; Ceccato, P. Comparison of CMORPH and TRMM-3B42 over the mountainous regions of Africa and South America. In *Satellite Rainfall Applications for Surface Hydrology*; Gebremichael, M., Hossain, F., Eds.; Springer: London, UK, 2010; pp. 193–204.
- 14. 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]
- Haile, A.; Habib, E.; Rientjes, T. Evaluation of the Climate Prediction Center (CPC) morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile River. *Hydrol. Process.* 2013, 27, 1829–1839. [CrossRef]
- 16. Ebert, E.E.; Janowiak, J.E.; Kidd, C. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteor. Soc.* **2007**, *88*, 47–64. [CrossRef]
- 17. Tian, Y.; Peters-Lidard, C. A global map of uncertainties in satellite-based precipitation measurements. *Geophys. Res. Lett.* **2010**. [CrossRef]
- 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. [CrossRef]
- 19. 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.* **2011**. [CrossRef]
- 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]
- 21. 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]
- 22. Derin, Y.; Yilmaz, K. Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. *J. Hydrometeorol.* **2014**, *15*, 1498–1516. [CrossRef]
- 23. Hirpa, F.A.; Gebremichael, M. Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *J. Appl. Meteorol. Climatol.* **2010**, *49*, 1044–1051. [CrossRef]
- 24. Hong, Y.; Gochis, D.; Cheng, J.; 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]
- 25. Kubota, T.; Ushio, T.; Shige, S.; Kida, S.; Kachi, M.; Okamoto, K. 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]
- 26. Demaria, E.M.; Serrat-Capdevila, A. PART III: Validation of remote sensing estimated meteorological variables. In *Earth Observations for Water Resources Management: Current Use and Future Opportunities for the Water Sector*; Luis, E., García, D.R., Marcus, W., Inge, P., Eds.; The Water Partnership Program—The World Bank: Washington, DC, USA, 2016; pp. 167–194.
- 27. Serrat-Capdevila, A.; Valdes, J.B.; Stakhiv, E.Z. Water Management Applications for Satellite precipitation products: Synthesis and Recommendations. *J. Am. Water Resour. Assoc.* **2014**. [CrossRef]
- 28. Adeyewa, Z.D.; Nakamura, K. Validation of TRMM radar rainfall data over major climatic regions in Africa. *J. Appl. Meteorol.* **2003**, *42*, 331–347. [CrossRef]

- 29. Romilly, T.G.; Gebremichael, M. Evaluation of satellite rainfall estimates over Ethiopian river basins. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1505–1514. [CrossRef]
- 30. 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. Hazard.* **2009**. [CrossRef]
- 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]
- 32. Casse, C.; Gosset, M.; Peugeot, C.; Pedinotti, V.; Boone, A.; Tanimoun, B.A.; Decharme, B. Potential of satellite rainfall products to predict Niger River flood events in Niamey. *Atmos. Res.* **2015**, *163*, 162–176. [CrossRef]
- 33. Ringard, J.; Becker, M.; Seyler, F.; Linguet, L. Temporal and spatial assessment of four satellite rainfall estimates over French Guiana and North Brazil. *Remote Sens.* **2015**. [CrossRef]
- 34. 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]
- 35. Ali, A.; Lebel, T.; Amani, A. Rainfall estimation in the Sahel. Part I: Error function. *J. Appl. Meteorol.* **2005**, 44, 1691–1706. [CrossRef]
- 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]
- Gruber, A.; Levizanni, V. Assessment of Global Precipitation Products: A Project of the World Climate Research Programme Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel; Report WCRP-128, WMO/TD-No. 143; World Climate Research Program: Geneva, Switzerland, 2008.
- 38. Huffman, G.; 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. Hydrometall.* **2007**. [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]
- 40. Hsu, K.L.; Gao, X.; Sorooshian, S.; Gupta, H.V. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.* **1997**, *36*, 1176–1190. [CrossRef]
- 41. Kottek, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World Map of the Köppen-Geiger climate classification updated. *Meteorol. Festschr.* 2006, *15*, 259–263. [CrossRef]
- Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* 2009, 377, 80–91. [CrossRef]
- 43. Winter, C.L.; Nychka, D. Forecasting skill of model averages. *Stoch. Environ. Res. Risk Assess.* 2010, 24, 633–638. [CrossRef]
- 44. Artan, G.; Gadain, H.; Smith, J.L.; Asante, K.; Bandaragoda, C.J.; Verdin, J. Adequacy of satellite derived rainfall data for stream flow modelling. *Nat. Hazard.* **2007**, *43*, 167–185. [CrossRef]
- 45. Gebregiorgis, A.S.; Hossain, F. Making satellite precipitation data work for the developing world. *IEEE Mag. Trans. Geosci. Remote Sens.* **2014**. [CrossRef]



© 2016 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/).