Validation of Satellite-Derived Rainfall Products over the Sahel

Nick Novella[1] and Wassila Thiaw[2]

(1) Wyle Information Systems / CPC / NOAA (2) CPC / NOAA
5200 Auth Road. Camp Springs, MD 20746 USA

Abstract

In this study, the primary goal was to evaluate the performance of the NOAA-CPC Rainfall Estimator (RFE), as well as a number of other high-resolution precipitation products over a specified region in the Sahel. Independent rain gauge data were collected from over 100 local stations in Mali, Senegal, and Burkina Faso during the summer season of 2008. After aggregation and interpolation, this data was specifically used to diagnose systematic differences between in-situ based rainfall and satellite-derived rainfall using an extensive selection of validation metrics. While previous literature describes weaknesses found in the CPC RFE precipitation product in other parts of Africa, this study determined that the CPC RFE performed reasonably well in the estimation and monitoring of Sahel rainfall in 2008.

Introduction

In recent decade, the advancement of satellite remote sensing data and applications has allowed the generation of numerous precipitation estimator products. Such products have consisted of uniquely-defined algorithms that merge rain gauge data with remotely sensed infra-red data from geostationary satellites, and passive microwave and active radar data from polar-orbiting satellites. In 1998, the Climate Prediction Center (CPC) developed and maintained a satellite rainfall estimator known as the RFE. The RFE generates gridded, daily precipitation estimates at a high 0.1˚ X 0.1˚ resolution centered over the African continent. The RFE algorithm consists of blending daily, automated rain gauge data (GTS), with geostationary infra-red (IR) and polar-orbiting passive microwave (PW) satellite data based on the methods of Xie and Arkin (1996). The primary motivation for creating the RFE was to allow a more accurate monitoring of regional and large-scale climatic and hydro-meteorological trends for the U.S. Agency of International Development, and the Famine Early Warning Systems Network (USAID / FEWS-NET) humanitarian aid programs. Since its inception, the RFE product continues to satisfy the demands of both operational analyses and research driven tasks here at CPC.

In recent studies, several merged satellite-gauge precipitation products have been subject to validation examinations to ascertain the accuracy of rainfall estimates on various space and timescales within the African continent. In these satellite-derived precipitation product inter-comparisons, literature has described a number of performance issues associated with the RFE product. In an East Africa validation study, Dinku et al. (2007) found that RFE rainfall estimates did not compare well with in-situ rain gauge data over a study region in Ethiopia. Specifically, he noted that topography plays a significant role in satellite rainfall estimation due to an algorithm's inability to detect oregraphically induced rainfall. While this weakness does not particularly impede an estimator's ability to detect the daily occurrence of precipitation, it does however lead to considerable misestimates relative to the magnitude of rainfall on a daily basis.

In this project, the authors wanted to expand upon previous RFE examinations by evaluating the performance of the RFE, as well as, other high resolution satellite rainfall estimators with rain gauge data in West Africa. The goal was to collect a number of in-situ rain gauges that were independent of the gauge data regularly ingested by the RFE on a daily basis to enable a true-reference and unbiased validation. This paper presents the validation metrics selected, and discusses the performance of the RFE and other rainfall estimators relative to in-situ gauge measurements during the summer rains season of 2008. It is intended that the results of this validation will promote confidence in the RFE's ability to detect rainfall over Africa, and inspire subsequent validation exercises for other remote regions in Africa.

Data and Methods

The in-situ rain gauge data used in the validation study was provided by African visitors at the Africa Desk, CPC. Gauge data from 133 stations were collected from the countries of Mali, Senegal and Burkina Faso in West Africa. This data is independent from the gauge data ingested by the satellite rainfall estimators selected in this study. All gauge data were thoroughly checked for reliability and consistency. Rain gauge measurements consisted of accumulated 24-hour rainfall totals from 06Z to 06Z and cover the period from June – September in 2008. Figure 1 illustrates the study region and spatial distribution of the gauge sites used in this validation.

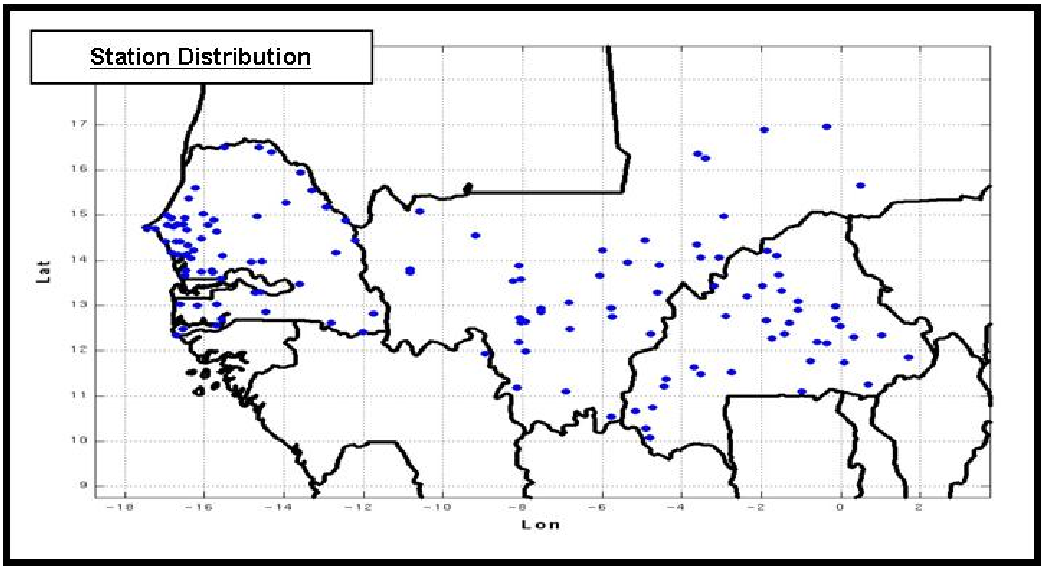

Figure 1: Spatial distribution of rain gauge stations used to validate satellite rainfall

Given the temporal and spatial characteristics of the gauge data, it was germane that the satellite rainfall estimators selected in this study were high-resolution products. Table 1 provides a summary of the high-resolution precipitation products (HRPP's) used in this validation. The authors acknowledge the CPC Unified Gauge dataset is not a satellite-derived product; however there was particular interest to evaluate its in-situ compatibility with the independent gauge data, and determine any distinguishable features associated with the input satellite data for all the other precipitation products.

HRPP	**Original Spatial / Temporal Resolution**	**Developed by:**
RFE	0.1 / Daily	CPC
ARC	0.1 / Daily	CPC
TRMM 3B42-RT	0.25 / 3-Hourly	NASA
CMORPH	0.25 / 3-Hourly	CPC
PERSIANN	0.25 / 3-Hourly	University of Arizona, US
TAMSAT	0.05 / Dekadal (10-day totals)	University of Reading, UK
CPC Unified Gauge	0.5 / Daily	CPC

Table 1: Summary of high-resolution precipitation products (HRPP's) selected in this study

The RFE (Hermann et al., 1997, Xie and Arkin, 1996) and ARC (Love et al. 2004, Novella et al. 2009) precipitation products are at a daily, 0.1° resolution covering the African continent. Both these products were developed by CPC and function by the same governing algorithm (Xie and Arkin, 1996). However, the primary differences between the RFE and the ARC lie in the usage of geostationary infra-red (IR) and polar-orbiting passive microwave (PM) data. Specifically, while the

RFE ingests half-hourly IR data and is merged with SSM/I and AMSU-B microwave data, the ARC product uses only three-hourly IR data and does not include the SSM/I and AMSU-B microwave inputs. The advantage of decreased IR sampling and excluded PM data in the ARC product was that it specifically allowed for the construction of a long-term, homogeneous precipitation dataset suitable for climatological analysis. In using only gauge and IR data that is consistent over time (Love et al. 2004, Novella et al. 2009) showed that the ARC method maintains high correlation with ground observations. Furthermore, the ARC compliments the RFE by capturing the inter-annual variability as well as large-scale patterns of African rainfall; however it was found that the ARC suffers slightly in the estimation of intensely convective precipitation on daily basis due to the exclusion of PM inputs.

The TRMM 3B42-RT (Huffman et al., 2003) precipitation product is a 3-hourly gridded rainfall analysis, at a 0.25˚ resolution over the globe. The key inputs for 3B42-RT are IR data from geostationary satellites, and polar-orbiting sensors that use PM and precipitation radar reflectivity (PR) data. Because this product is produced on a real-time basis (with a minor lag of 6 hours), gauge data are not included in this particular product version. Merging methodology of the 3B42-RT product is described in (Huffman et al. 2003, Huffman et al. 2007). The CMORPH precipitation product (Joyce et al., 2004) is another real-time “satellite-only” estimator that uses IR and PM data. Unlike most satellite rainfall products, CMORPH is unique because rainfall estimates are generated by using motion vectors derived from half-hourly IR data, which is then used to provide displacement vectors for every observation interval in the PM data. This “morphing” process yields the accuracy of PM rainfall estimates with the higher temporal and spatial resolution of the IR data. The CMORPH data used in this study consisted of aggregated 3-hourly rainfall estimates at 0.25˚ resolution. The PERSIANN precipitation product (Sorooshian et al., 2005) also uses only IR and PM data, but implements an artificial neural network technique to generate a high-resolution gridded analysis. This technique, which produces 3-hourly 0.25˚ global rainfall estimates, operates in two real-time modes (i.e simulation and update). In the first mode, rainfall estimates are based on a previously calibrated neural network mapping function, whereas the second mode uses instantaneous PM data, when available, to adjust the parameters employed by the mapping function.

Similar to the RFE and ARC, the TAMSAT precipitation product (Grimes et al., 1999, Thorne et al., 2001) exclusively covers the African continent; however, rainfall estimates are only available as dekadal (10-day) totals at a 0.05˚ resolution. Methodologically, the TAMSAT product uses only geostationary IR data onboard Meteosat satellites. Where (Arkin and Meisner, 1987) found that using a fixed brightness temperature threshold to calculate rainrates in tropical rain producing clouds (i.e. GPI), the temperature thresholds used in TAMSAT vary spatially and temporally based on calibration with in-situ rain gauge data. In essence, TAMSAT evaluates the length of time that an IR pixel exceeds a temperature threshold is used to create cold cloud duration (CCD) images, which is then linearly related to rainfall accumulated for the dekadal period. Lastly, the CPC Unified gauge dataset is a relatively new precipitation product developed by CPC (Chen et al., 2008). The key advantage in using this new gauge-based dataset is that it exhibits high correlation and reduced bias with independent rain gauge measurements, when compared to other objectively defined ground-based precipitation analyses. The Optimal Interpolation (OI) technique (Gandin, 1965) employed in this CPC Unified gauge dataset was found to possess the highest performance statistics over regions characterized by both dense and sparsely distributed stations. Given the gauge distribution shown in Figure 1, the authors felt this particular gauge dataset was ideal for this study.

For consistency with daily gauge measurements, all High Resolution Precipitation Products (HRPP’s) were temporally averaged from 06Z-06Z (except TAMSAT) per each 24-hour daily estimate, and summed into dekadal rainfall totals for the period covering June to September, 2008. The spatial domain designated for this validation covers -18.75˚ W to 3.75˚ E and 8.75˚ N to 18.75˚ N on a 0.5˚ resolution grid. To ensure a uniform validation with independent gauge data, all HRPP’s were spatially averaged from their original resolution to match the resolution of the unified 0.5˚ grid.

However, the common challenge when comparing gauge (point-based) data and gridded (area-based) data is how to resolve their spatial incongruence. Several works in literature have attempted to develop systematic methods to minimize the error between an in-situ point value and an interpolated value on a regularly spaced grid. Previous rainfall validation studies have found that optimal methods require an algorithm that is not only dependent on the observed point value located within an areal grid, but also dependent on the distribution, stability, and magnitude of neighboring data points over various space and time scales. To account for these previous findings in converting the independent point data to gridded, areal values, the validation of each HRPP in this study consisted of using three different approaches to generate an areal 0.5° gauge estimate. These independent areal gauge estimates were then paired with the HRPP rainfall estimates between two collocated grids for each dekad. For the first approach, the areal averaged gauge value consisted of a simple averaging of rain gauge point totals inside a 0.5° grid, but only for grids that contained at least one station. This condition allowed only 88 grids out of the 966 total number of grids in the study region. For all 12 dekads from June to September, this equated to 1056 data pairs for validation. The second approach shared the same grid requisite of the first approach; however a geo-statistical technique was additionally implemented. An error function based on the works of (Ali et al., 2005) was developed to analytically yield the error associated with the point data from the gauges and the respective areal estimate. According to (Ali et al., 2005), this error function (Expression 1) is based on the time-scaling properties of the rain fields and was successfully validated at the mesoscale as part of a validation study using Sahel rainfall data. It is expressed as:

$$e(A, N_g, K_T, P_T) = \frac{1.05}{\sqrt{N_g}\sqrt{K_T}}\left(\frac{P_T}{K_T}\right)^{-0.2} \times \left[0.25 + 0.11\log\left(\frac{A}{N_g}\right)\right] + 0.03 \qquad \textbf{(1)}$$

Expression 1: Error Function used to estimate areal rainfall from rain gauges. Yields error in percent for area considered (Ali et al, 2005).

A is the area (km^2) for the areal estimate, $\mathbf{N_g}$ is the number of gauges located within the area, and $\mathbf{K_T}$ and $\mathbf{P_T}$ are the number of rain events and rainfall total for dekad, respectively. The threshold criterion was designated at an error level of less than 35% for all 0.5° gridpoints in the domain. Thus, this error threshold used in the second approach placed further constraint on the total number of grids, resulting in a lesser number of data pairs in the validation. Based on the recent findings associated with optimal interpolation (Gandin, 1965), the third approach used this technique on the independent gauge data. Similar to the first and second approaches, only areal gauge estimates containing at least one station in the 0.5° grid were used. Results from each of these approaches were later examined to determine any differences between them.

For each approach, there were a number of measures/indices selected to examine differences between the independent gauge data and the HRPP estimates. For the validation itself, the authors considered the independent areal gauge estimates (derived from the three approaches) as the observed “ground-truth” for all validation measures. These measures were divided into three different categories: dichotomous, continuous, and precision. Dichotomous measures were used to diagnose the HRPP’s ability to detect daily precipitation events, where rainfall events are defined by rain (>= 1.0 mm) / no rain (<1.0 mm) thresholds on a daily basis. In this scheme, a contingency table is used that aggregates rain / no rain occurrences to collect the number of hits, misses, correct negatives, and false alarms between the paired precipitation data. Expressions 2-5 were computed from the contingency table to describe particular aspects of the HRPP’s performance relative to frequency of detection.

$$POD = \frac{hits}{hits + misses} \qquad \textbf{(2)}$$

$$POFD = \frac{false\ alarms}{correct\ negatives + false\ alarms} \qquad \textbf{(3)}$$

$$HK = \frac{hits}{hits + misses} - \frac{false\ alarms}{false\ alarms + correct\ negatives} \qquad \textbf{(4)}$$

$$OR = \frac{hits \cdot correct\ negatives}{misses \cdot false\ alarms} = \frac{\left(\frac{POD}{1-POD}\right)}{\left(\frac{POFD}{1-POFD}\right)} \qquad \textbf{(5)}$$

Expressions2-5: Dichotomous measures based on contingency table of rain/no rain occurrences between HRPP and independent gauge data.

The probability of detection (**POD**) measures the percentage of real precipitation events that were correctly detected by the HRPP. Conversely, the probability of false detection (**POFD**) measures the number of no rain events that were incorrectly detected as real events by the HRPP. The Hanssan and Kuiper's discriminate (**HK**) measures the HRPP's ability to separate real versus non-real rain events, which essentially combines the first two POD and POFD measures. Higher scores are yielded for correctly detecting frequent events, which is particularly useful for regularly occurring rains during summer in West Africa. The odds ratio (**OR / ODDS**) is the probability of the HRPP correctly identifying a real precipitation to probability of the HRPP incorrectly identifying a real precipitation event. This measure yields higher scores for infrequent rain events. All dichotomous scores are illustrated in the form of bar charts for each HRPP.

A small number of continuous measures were selected to determine the HRPP's ability to accurately identify the magnitude of precipitation, as well as, to diagnose systematic errors associated with the HRPP on a dekadal basis. In essence, continuous methods depict how the values of HRPP's differ from those observed in the independent gauge data. They are shown in the expressions 6-8, where **F** and **O** represent the HRPP and independent gauge totals for the dekad, respectively:

$$r = \frac{\sum(F-\bar{F})(O-\bar{O})}{\sqrt{\sum(F-\bar{F})^2}\sqrt{\sum(O-\bar{O})^2}} \quad \textbf{(6)} \qquad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i-O_i)^2} \quad \textbf{(7)} \qquad Bias = \frac{\frac{1}{N}\sum_{i=1}^{N}F_i}{\frac{1}{N}\sum_{i=1}^{N}O_i} \quad \textbf{(8)}$$

Expressions6-8: Continuous measures to depict dekadal differences in magnitude between HRPP and independent gauge data.

Although correlation (**r**) is the most direct and linear measurement in terms of how well the HRPP and gauge data correspond for validation. Two weaknesses associated with correlation are that extreme HRPP errors may be compensated by nonconcurring extreme gauge values, and that it does not account for bias. To mitigate this, root mean square error (**RMSE**) and multiplicative bias (**BIAS**) measures were additionally used. RMSE is favorable for its ability to penalize the HRPP for large errors despite the direction of deviation from the gauge value (i.e. negates compensating errors). BIAS simply compares the average magnitudes between the HRPP and gauge data in the form of a ratio, which is then used to infer any systematic error associated with the HRPP. All continuous scores are illustrated in a bar chart to show the relationship between the independent gauge estimates and the respective HRPP estimate for all 12 dekads.

Precision performance measures were used to observe any systematic changes between the HRPP and gauge estimates over varying magnitudes of binned rainfall. For this particularly set of validation, only gauge estimates that employed the error function from approach 2 were used. In essence, precision validation determined the relative departure of the HRPP value from the observed independent gauge dekadal value. If HRPP value were within ± 2 mm of the gauge estimate, it was assigned a "near-equal" score. If a HRPP estimate was within ± 10 mm of the gauge estimate, it was assigned a "slightly over / under estimate "score. If the HRPP estimate exceeded 10mm in either direction, it was assigned an "over / under estimate score". Collectively, these scores were then divided into bin categories based on the magnitude of the observed gauge precipitation. The three bins selected for the precision examination were for all dekadal gauge values <50 mm, >=50 and <100mm, and >100mm. The results of the precision validation are illustrated in the form of pie charts.

Results and Discussion

The following section discusses the validation results from the 3 approaches methods used for all seven HRPP's. For the dichotomous validation results, the CMORPH and PERSIANN products scored the highest relative to their ability to detect precipitation events (POD). However, POFD scores suggest that PERSIANN incorrectly estimated rainfall more time where there was no rainfall on the ground according to the independent gauge data. All other products showed an even distribution of detection scores for all three approaches. HK scores indicate that the RFE and CMORPH products performed the best relative to their ability to separate rain versus non-rain events according to the gauge data. This is also reflected in the OR/ODDS score, suggesting that both the RFE and CMORPH are most likely to correctly identify rain events, and least likely to make false alarms on a daily basis. These scores are illustrated in figure 2.

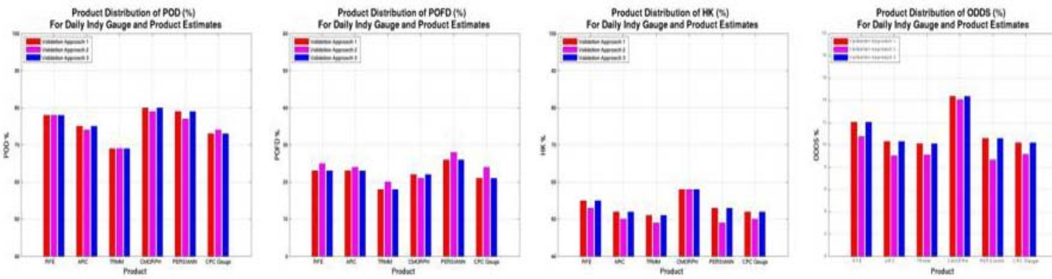

Figure 2: Series of bar-plots depicting dichotomous scores between HRPP's and independent gauge rainfall for all three validation approaches.

For continuous validation, all three approaches showed that the RFE and ARC products performed the best relative to having the highest correlation, with considerably low BIAS and RMSE scores. Conversely, TRMM and PERSIANN yielded the lowest correlations with the independent gauge data for all 12 dekads. The RFE, ARC, TAMSAT and CPC Unified gauge data all yielded the lowest RMSE amongst all other HRPP's, suggesting that TRMM, CMORPH and PERSIANN estimates are more likely to possess larger differences with the observed in-situ data over the Sahel. Similarly, the RFE, ARC, TAMSAT, and CPC Unified gauge rainfall estimates consistently showed to have the lowest BIAS, suggesting less systematic error associated with these products. These scores are reflected in figure 3.

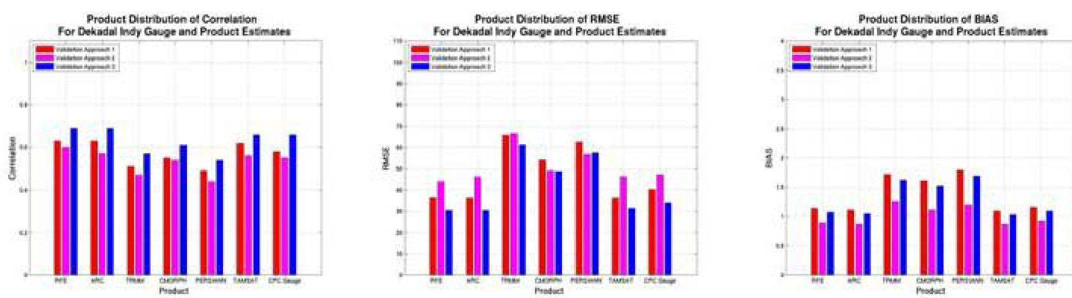

Figure 3: Series of bar-plots depicting continuous scores between HRPP's and independent gauge rainfall for all three validation approaches.

Precision performance scores indicate that for all independent rain gauge totals less than 50 mm, the RFE, ARC, and CPC Unified products maintained the lowest percentages of over / under estimates for all dekads. The authors note that a majority of over-estimates between all products generally occurred when there was no rain reported on the ground. For all independent rain gauge totals greater than 50 mm, and less than 100mm, the ARC was observed to have fewest over / under estimates compared to the other products, with 36% (31%) of observations over (under) estimating by at least 10 mm for all dekads. For all independent rain gauge totals greater than 100mm, it was evident that all products have the tendency to underestimate high rainfall reported on the ground. However, the ARC and PERSIANN products were found to have the highest percentage of observations closest to ground reports. The percentages of estimates relative to the independent gauge data are illustrated in figure 4

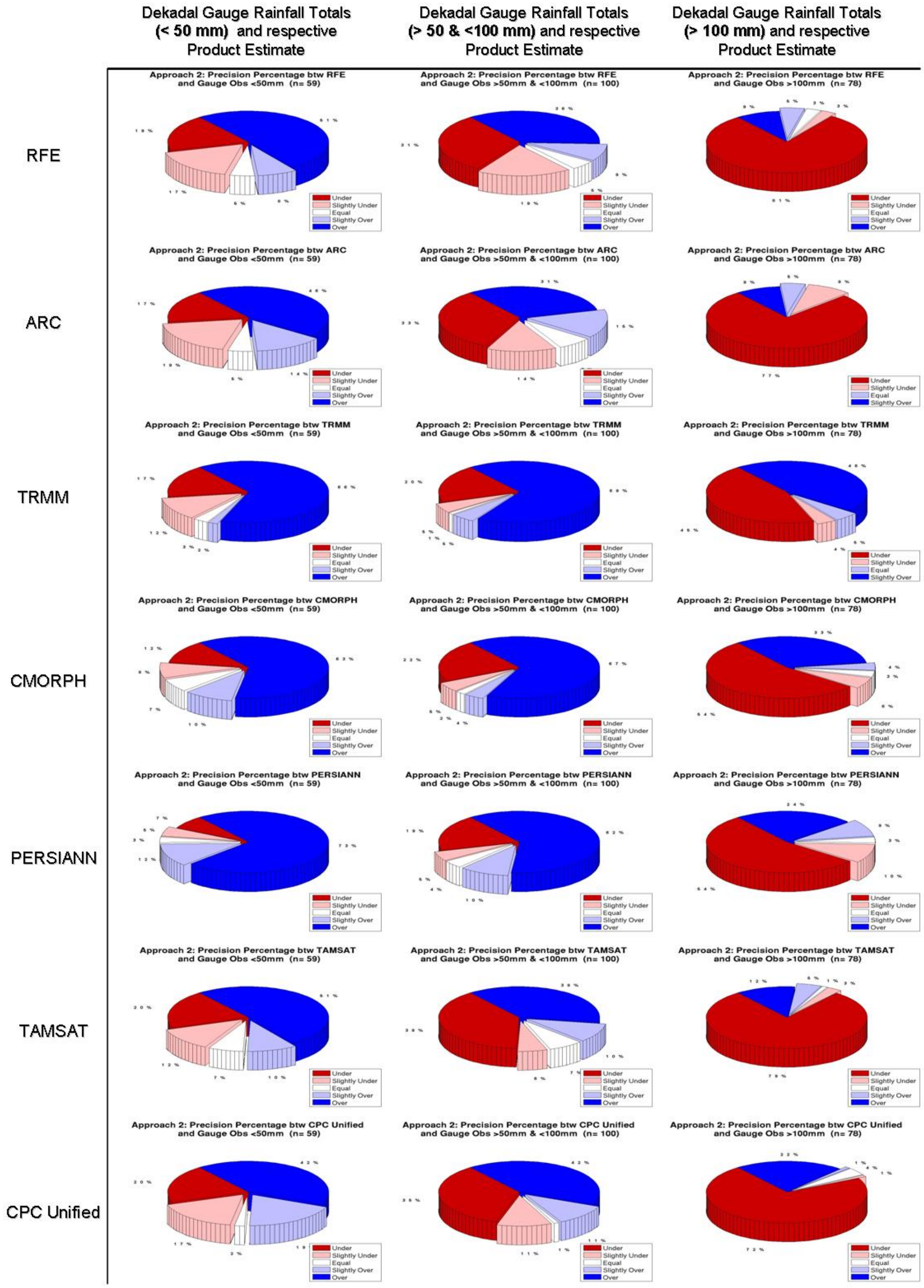

Figure 4: Series of pie-charts depicting precision scores between HRPP's and independent gauge rainfall.

Conclusions

After the various types of validation conducted in this study, it was found that the RFE and ARC performed the best overall with the independent gauge data reported during the summer of 2008. Specifically, performance scores remained consistently high in terms of detecting a rain event itself, and identifying the magnitude of rainfall for weak and extreme rainfall totals. A major caveat of this study is that this validation was not performed over high terrain and complex topography where the RFE was shown to suffer the most in literature. Future validation examinations aim to accomplish this, in order to gain a closer look at areas where the RFE performance is potentially weak. Despite consistency in performing well in the CPC Unified gauge dataset, the absence of satellite input in this product did not appear to be considerably more or less advantageous over the other satellite-based products. The greatest weakness seen with the gauge-only data is its potential for poor accuracy associated with high rainfall totals according to the independent gauge data. This is likely the result of the parameterization of the OI technique used. Lastly, the authors cannot discern which approach is the best method of validation, as all exhibited consistently similar results. Although the inclusion of the error function suggests a more credible areal estimate derived from the independent gauge measurements, a significantly lesser number of cases associated with approach 2 did not ensure statistically significant results. These points demonstrate the need for additional data that includes other years so that an independent rain gauge climatology may be developed, as well as, the necessity to apply these validation methods elsewhere across other African regions.

References

Ali A, Lebel T, and Amani A. (2005) Rainfall Estimation in the Sahel. Part I: Error Function. Journal of Applied Meteorology, **44**, pp 1691-1706

Ali A, Amani A, Diedhiou A, and Lebel T. (2005) Rainfall Estimation in the Sahel. Part II: Evaluation of Rain Gauge Networks in the CILSS countries and Objective Inter-comparison of Rainfall Products. Journal of Applied Meteorology, **44**, pp 1707-1722

Arkin PA, and Meisner BN. (1987) The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. Monthly Weather Review, **115**, pp 51-74. DOI: 10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2

Chen M, Shi W, Xie P, Silva V, Kousky VE, Higgins WR, Janowiak J. (2008) Assessing objective techniques for gauge-based analysis of global daily precipitation. Journal of Geophysical Research, **113**, D04110. DOI: 10.1029/2007JD009132

Dinku T, Ceccato P, Grover-Kopec E, Lemma M, Connor SJ, Ropelewski CF. (2007) Validation of satellite rainfall products over East Africa's complex topography. International Journal of Remote Sensing, 28, pp 1503-1526. DOI: 10.1080/01431160600954688

Gandin, LS. (1965) Objective Analysis of meteorological fields. Israel Program for Scientific Translations, 242 pp.

Grimes DF, Pardo E, and Bonifacio R. (1999) Optimal areal rainfall estimation using rain gauges and satellite data. Journal of Hydrology, **222**, pp 93-108

Herman A, Kumar V, Arkin PA, Kousky JV. 1997. Objectively determined 10-day African rainfall estimates created for famine early warning. *International Journal of Remote Sensing* **18**: 2147-2159. DOI: 10.1080/014311697217800

Huffman GJ, Adler RF, Stocker EF, Bolvin DT, and Nelkin EJ. (2003) Analysis of TRMM 3-Hourly multi-satellite precipitation estimates computed in both real and post-real time. Combined Preprints CD-ROM, 83rd AMS Annual Meeting Poster P4.11, In 12th Conference on Satellite Meteorology and Oceanography, 9-13 February 2003, Long Beach CA.

Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB. (2007) The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology, **8**, pp 38-88

Joyce RJ, Janowiak JE, Arkin PA, and Xie P. (2004) CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometerology, **5**, pp 487-503

Lebel T, and Amani A. (1999) Rainfall Estimation in the Sahel: What is the ground truth? Journal of Applied Meteorology, **38**, pp 555-568

Love TB, Kumar V, Xie P, Thiaw W. (2004) A 20-year daily Africa precipitation climatology using satellite and gauge data. In Proceedings of the 84th AMS Annual Meeting, P5.4. Conference on Applied Meteorology 11-15 January 2004, Seattle, WA.

Novella N, Thiaw W, Wolvovsky E, McCoy, Kumar V. (2009) A New 26-year African Rainfall Climatology. (In Press). International Journal of Climate.

Sooroshian S, Hsu K, Gao X, Gupta H, Imam B, Braithwaite D. (2005) Evaluation of PERSIANN System Satellite-Based Estimaes of Tropical Rainfall. Bulletin of the American Meteorological Society, **81**, pp 2035-2046

World Climate Research Programme. WWRP/WGNE Joint Working Group on Verification. Forecast Verification – Issues, Methods and FAQ. Australian Bureau of Meteorology. http://cawcr.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html

Xie P, Arkin PA. 1996. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. Journal of Climate, **9**, pp 840-858. DOI: 10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2