Article

Not peer-reviewed version

Spatiotemporal Evaluation of Multi-Source Precipitation Products in the Sudan Sahel: Evidence from White Nile State

Abdelbagi Yanes Fadlalmwlla Adam * , Zoltán Gribovszki , Péter Kalicz

Posted Date: 19 May 2026

doi: 10.20944/preprints202605.1175.v1

Keywords: CHIRPS; ERA5-Land; IMERG; precipitation evaluation; satellite rainfall estimates; Sudan Sahel; White Nile State; hydro-climatic monitoring; bias correction

Article

Spatiotemporal Evaluation of Multi-Source Precipitation Products in the Sudan Sahel: Evidence from White Nile State

Abdelbagi Yanes Fadlalmwlla Adam [1,2,*], Zoltán Gribovszki [1] and Péter Kalicz [1]

[1] Institute of Geomatics and Civil Engineering, University of Sopron, 9400 Sopron, Hungary
[2] Department of GIS and Cartography, University of Khartoum, P.O. Box 321, Khartoum 11115, Sudan
* Correspondence: abdelbagi.yanes@phd.uni-sopron.hu

Abstract

Accurate rainfall estimates are essential for managing water resources and planning for climate risks in semi-arid regions, yet long-term gauge networks in these environments are often extremely limited. In this study, we evaluate three widely used multi-source precipitation datasets; CHIRPS, IMERG, and ERA5-Land, against long-term observations from Ed Dueim and Kosti, the two main reference stations in White Nile State, central Sudan. The assessment covers monthly and annual scales across each product's available record (1952–2022) and uses a broad set of metrics, including Pearson and Spearman correlations, NSE, KGE, RMSE, MAE, percent bias, and categorical detection scores (POD, FAR, CSI). All three datasets capture the region's single-peak June–October monsoon pattern, but their accuracy differs sharply when it comes to rainfall amounts and year-to-year variability. CHIRPS performs best overall, with monthly NSE values around 0.77 and KGE between 0.79 and 0.88, along with a consistent dry bias of 5–13%—a predictable error that can be corrected operationally. IMERG shows strong monthly correlations but consistently overestimates rainfall by 25–42%, which leads to unreliable annual totals (NSE = −1.93 to −2.21). ERA5-Land performs worst across nearly all metrics, with monthly NSE near or below zero, annual NSE dropping to −15.34, and frequent false alarms during the dry season. Taken together, the evidence points to CHIRPS as the most reliable dataset for routine hydro-climatic monitoring in White Nile State, while IMERG and ERA5-Land may still be useful in more specialized or time-specific applications.

Keywords: CHIRPS; ERA5-Land; IMERG; precipitation evaluation; satellite rainfall estimates; Sudan Sahel; White Nile State; hydro-climatic monitoring; bias correction

1. Introduction

Rainfall is the main force shaping hydrological systems in semi-arid regions. It determines how much water is available, how crops perform, how ecosystems function, and how often communities face floods or droughts. In the Sudanese Sahel, where yearly rainfall ranges from only about 100 to 500 mm and arrives almost entirely during the short June–October monsoon (A. A. M. Salih et al., 2015). Even small errors in estimating precipitation can lead to major misunderstandings of water balance, drought timing, or runoff generation (Woods et al., 2023). Reliable, continuous, and spatially detailed rainfall monitoring isn't just a scientific goal here; it's a practical necessity for the people and economies that depend on these fragile landscapes.

The White Nile State of central Sudan exemplifies the broader challenge confronting hydro-climatic science across the Sahel the historical reliance on a sparse in-situ rain gauge network has produced persistent "data voids" in the spatial and temporal coverage of rainfall records (Funk et al., 2015). The gauge network at Ed Dueim and Kosti, while representing one of the longest continuous records available in the region (Adam et al., 2025). It provides only point-scale observations incapable of capturing the high spatial heterogeneity of convective rainfall events that characterize the semi-

 doi:10.20944/preprints202605.1175.v1

arid monsoon environment. The position of the Intertropical Convergence Zone (ITCZ) and the associated West African Monsoon system generate organized but spatially localized precipitation events that cannot be adequately represented by point measurements alone (Rauch et al., 2025).

To bridge this observational gap, Satellite-based Rainfall Estimates (SREs) and atmospheric reanalysis products have achieved widespread adoption across data-sparse regions. Three products in particular have become central to regional hydro-climatic research and operational monitoring; the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), which merges cold cloud duration satellite estimates with gauge observations and extends from 1981 to near-present at 0.05deg resolution (Funk et al., 2015); the Integrated Multi-satellitE Retrievals for GPM (IMERG), the successor to the Tropical Rainfall Measuring Mission (TRMM) product family, which provides half-hourly global estimates at 0.1deg resolution from 2000 onward (Huffman et al., 2020); and the ERA5-Land reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF), which offers a physically consistent estimate of land-surface conditions at hourly temporal resolution and 0.1deg spatial resolution extending back to 1950 (Muñoz-Sabater et al., 2021). Despite their widespread use, the performance of these products varies substantially across climatic settings, and independent regional validation against gauge observations is a prerequisite before their use in hydrological applications.

Previous work across sub-Saharan Africa and the wider Sahel has compared satellite-based rainfall estimates, often finding CHIRPS to be one of the strongest performers at the monthly scale, while also noting IMERG's tendency toward wet biases and the common underestimation issues in reanalysis products (Toté et al., 2015;Dinku et al., 2018); Satgé et al., 2020). But how these products behave under the specific conditions of central Sudan, a region with a very short, intense rainy season, a long dry season, and extremely sparse gauge coverage, remains only partly understood. Another gap is how their performance changes with the timescale. Studies in similar environments have shown that skill often drops sharply when moving from monthly to annual totals ((Teutschbein & Seibert, 2012; W. Salih et al., 2023) Yet this issue has received little attention in this part of Sudan.

This study fills those gaps by providing a detailed, multi-metric evaluation of CHIRPS, IMERG, and ERA5-Land using long-term gauge data from Ed Dueim and Kosti. We compare the products across monthly and annual scales, across seasons, and using both continuous statistics and categorical detection metrics, offering the most complete assessment to date for White Nile State. Our goals are fourfold: (1) to describe each product's climatological biases across the year; (2) to quantify statistical agreement and predictive skill at monthly and annual timescales; (3) to assess how well each product detects rainfall events in different seasons; and (4) to translate these findings into practical recommendations for product choice and bias correction in hydro-climatic applications.

 doi:10.20944/preprints202605.1175.v1

2. Study Area

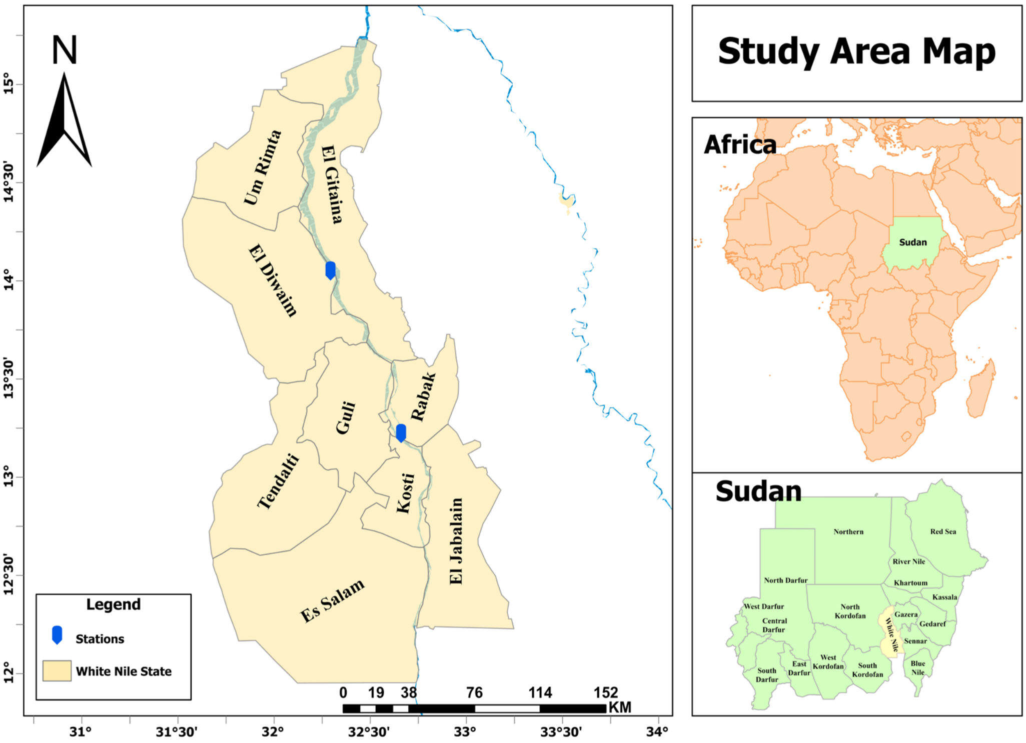

Figure 1. Study area map.

3. Data and Gridded Products

3.1. In-Situ Gauge Records

Meteorological data were obtained from the Sudanese Meteorological Authority (SMU) computer center of the Ministry of Irrigation and Water Resources (MIWR). The study area includes two stations, *Kosti* and *Ed Dueim*, as shown in Table 1, which provide rainfall and temperature data. The data cover monthly rainfall and temperature records spanning seven decades, from 1952 to 2022.

Table 1. Meteorological Stations in White Nile State, Sudan.

M. stations	Latitude (°N)	Longitude(°E)	Elevations/m	Rainfall Annual Mean	Mean Temperature
Kosti	13°.17'	32°.67'	386	431.26	28.70
Ed Dueim	13°.99'	32°.30'	378	266.03	29.22

3.2. Gridded Precipitation Products

CHIRPS (v2.0). The Climate Hazards Group Infrared Precipitation with Station data product integrates thermal infrared cold-cloud duration estimates from geostationary satellites with the CHPclim climatology and gauge observations using a modified inverse-distance-weighting interpolation scheme. CHIRPS is available globally between 50 °S and 50 °N at 0.05 ° spatial resolution and daily temporal resolution, with a continuous record from January 1981 to the present. Monthly aggregates were used in this study (Funk et al., 2015).

IMERG (Final Run, v06/v07). The Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM) mission provides merged estimates from passive microwave sensors, infrared-based retrievals, and monthly gauge analyses. IMERG Final Run products incorporate gauge adjustment and are available globally at 0.1° spatial resolution and 30-minute temporal resolution from June 2000 onward. For this study, we used the monthly aggregates from the IMERG Final Run product, covering the 2002–2022 period. (Huffman et al., 2020b).

ERA5-Land is a reanalysis dataset produced by ECMWF using the HTESSEL land-surface model, driven by ERA5 atmospheric fields. It has a spatial resolution of about 0.1° (~9 km) and provides hourly data. The record extends from January 1950 to the present and includes physically consistent estimates of land-surface variables, including total precipitation. For our analysis, we extracted monthly precipitation totals from the ERA5-Land grid cell corresponding to each gauge station for the full 1952–2022 evaluation period (Muñoz-Sabater et al., 2021).

4. Methods

4.1. Spatial Extraction and Temporal Aggregation

For each gridded dataset, we extracted the grid cell closest to each gauge station's location and converted the values to monthly totals for comparison with the station records. We chose nearest-cell extraction rather than spatial averaging to avoid adding extra smoothing that could hide local-scale biases. All products were evaluated at the monthly scale, and annual performance was assessed by summing monthly totals into calendar-year values.

4.2. Evaluation Period and Overlap

Each product was evaluated over the longest period of overlap available at each station: CHIRPS from 1981–2022 (41 years), IMERG from 2002–2022 (20 years), and ERA5-Land from 1952–2022 (71 years). All statistics were calculated within these respective periods. Seasonal analysis followed the region's standard hydro-climatic divisions winter (November–February), pre-monsoon (March–May), and the rainy season (June–October) (Elagib, 2010).

4.3. Continuous Performance Metrics

To assess how well each dataset matched the gauge observations, we used a set of complementary statistical metrics designed to capture different aspects of performance. Table 2 summarizes the formulas, value ranges, and optimal conditions for each metric.

Table 2. Statistical and categorical metrics are used to evaluate satellite precipitation products and reanalysis data against in-situ gauge observations.

Metric	Equation	Value Range	Optimal Value	Purpose	Reference
Spearman's Correlation	$\rho = \frac{cov(R_s\, R_o)}{(\sigma R_s\, \sigma R_o)}$	-1 to +1	+1	Analyzes the monotonic relationship between satellite data and gauge data.	(Spearman, 1961)
Bias	$\text{Bias} = \sum (s - o)$	(-∞, +∞)	0	Mean difference between satellite and observed values	(Willmott et al., 1985)
Percent Bias	$\text{PBIAS} = \frac{\sum(s-o)}{\sum o} 100$	(-∞, +∞)	0%	Indicates systematic over- or underestimation	(Gupta et al., 1999)
Mean Absolute Error	$\text{MAE} = \frac{\sum \lvert s - o \rvert}{N}$	> =0	0	Average magnitude of errors	Willmott & Matsuura (2005)
Root Mean Square Error	$\text{RMSE} = \frac{\sqrt{s - o}}{N}$	>=0	0	Measure the overall error between the satellite observations and the actual observations.	Willmott & Matsuura (2005

Nash-Sutcliffe Efficiency	$\text{NSE} = \frac{\sum(s-o)2}{\sum(o-\mu_o)2}$	(-∞,1)	1	Measures predictive skill	Nash & Sutcliffe (1970)
Kling-Gupta Efficiency	$\text{KGE} = \sqrt{(\{r-1)\}^2(\beta-1)^2(\gamma-1)^2}$	(-∞,1)	1	Integrates correlation, bias, and variability	Gupta et al. (2009)
Probability of Detection	POD = H / (H + M)	0 - 1	1	Detection accuracy is the degree to which a model correctly detects events that occurred.	Wilks (2011)
False Alarm Ratio	FAR = F / (H + F)	0 - 1	0	Measures of false rainfall detections	Wilks (2011)
Critical Success Index	CSI = H / (H + M + F)	0 - 1	1	CSI evaluates the overall accuracy of event detection,	Jolliffe & Stephenson (2012)

To assess the fidelity with which satellite-derived estimates reproduce ground-observed rainfall, we used the Pearson correlation coefficient (r) and the coefficient of determination (R^2). These metrics quantify the strength of the straight-line relationship between the two datasets. However, because rainfall distributions in semi-arid environments are typically skewed, exhibit high temporal variability, and contain frequent zero values, we additionally computed Spearman's rank correlation (ϱ), a non-parametric measure of monotonic association that assesses whether higher observed values consistently correspond to higher satellite estimates regardless of the precise magnitude of those values and is therefore less sensitive to outliers and distributional asymmetry than Pearson's r.

Predictive skill was evaluated using the Nash–Sutcliffe Efficiency (Nash & Sutcliffe, 1970), which benchmarks product performance against the long-term observed mean as a constant predictor. An NSE of 1 indicates perfect agreement; a value of 0 indicates that the product performs no better than the observed mean as a predictor of individual values; and negative values indicate that the observed mean is a more reliable predictor than the satellite product itself. To complement NSE, we also computed the Kling–Gupta Efficiency (Gupta et al., 2009), which addresses the known tendency of NSE to disproportionately weight large rainfall events—a direct consequence of its squared-error formulation by decomposing total model error into three dimensionless component ratios; the linear correlation coefficient, the bias ratio (mean satellite estimate to mean observed), and the variability ratio (satellite to observed coefficient of variation). This decomposition enables

systematic diagnosis of whether poor efficiency arises from timing errors, magnitude offsets, or variability mismatches.

Systematic and absolute errors were characterized through total accumulated bias (mm) and percent bias (PBIAS), which identify persistent tendencies toward over- or underestimation across the evaluation period. The typical magnitude of errors was quantified using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Because RMSE squares each error before averaging and then takes the square root, it gives much more weight to large mistakes than to small ones. This makes it especially sensitive to how well a satellite product performs during heavy rainfall events, where big deviations tend to dominate the error structure.

5. Results

5.1. Seasonal Rainfall Climatology (2002-2022)

The monthly rainfall climatology of Ed Dueim and Kosti over the 2002–2022 period exhibits a sharply defined unimodal annual cycle characteristic of the semi-arid Sudan Sahel zone (Figure 2). Observed rainfall at both stations remains effectively zero (< 1 mm/month) from November through April, with meaningful accumulations commencing in June, intensifying to a single peak in August, approximately 105 mm/month at Ed Dueim and 126 mm/month at Kosti before declining rapidly through September and returning to negligible levels by October. This seasonal structure is governed by the northward penetration of the West African Monsoon system, encompassing organized southwesterly moisture flux, the position of the continental monsoon trough, and the activity of mesoscale convective systems rather than ITCZ displacement alone (Bell et al., 2006). All three gridded products, CHIRPS, IMERG, and ERA5-Land, accurately reproduce the unimodal structure and the maximum August rainfall, thereby confirming adequate representation of large-scale monsoon timing. However, the products vary considerably in quantity. IMERG shows a clear wet bias at both stations, overestimating maximum August precipitation by about 25 mm (~135 mm) at Ed Dueim and by 45 mm (~170 mm) at Kosti, representing relative overestimates of about 22% and 36%, respectively. In contrast, ERA5-Land shows a consistent and systematic dry bias throughout the rainy season, with maximum monthly values of only ~ 67 mm at both stations, underestimating the observed August total by about 40%. CHIRPS provides the closest agreement with observations, tracking the August peak and seasonal recession most faithfully at both locations. These contrasting bias instructions emphasize that agreement in seasonal times is a necessary but wholly insufficient criterion for product evaluation, and that rigorous bias characterization over the entire annual cycle is still necessary before adopting a gridded product for hydrologic or water resource applications in this region.

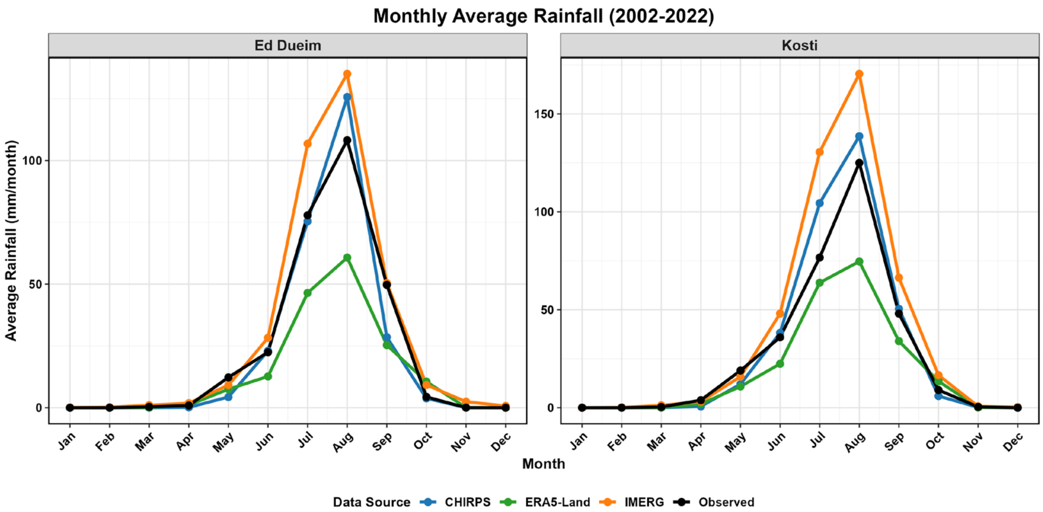

Figure 2. Monthly mean rainfall (2002–2022) for Ed Dueim and Kosti derived from CHIRPS, ERA5-Land, IMERG, and observations. All datasets show a pronounced June–October rainy season with peak rainfall in August.

5.2. Annual Average Rainfall and Interannual Variability (2002-2022)

The consistent ranking of products observed in the monthly climatology is preserved in mean annual totals (Figure 3). At Ed Dueim, the observed mean annual rainfall is 276.1 mm/year. CHIRPS produces the closest estimate at 260.9 mm/year (-5.5%), followed by IMERG at 345.2 mm/year (+25.0%), while ERA5-Land records the lowest total at 164.5 mm/year (-40.4%). At Kosti, the observed mean is 318.3 mm/year; CHIRPS again performs closest at 350.5 mm/year (+10.1%), IMERG overestimates at 453.3 mm/year (+42.4%), and ERA5-Land records 221.3 mm/year (-30.5%). The product hierarchy is consistent at both sites: IMERG > Observed ~ CHIRPS > ERA5-Land. All three datasets capture the north–south rainfall gradient, showing that the broad spatial pattern is well represented across products. The error bars in Figure 2 also make it clear that each dataset reflects year-to-year variability, though IMERG and ERA5-Land show a wider spread than both the gauge record and CHIRPS. That extra scatter likely comes from retrieval noise, model uncertainty, and how each product handles extreme events.

 doi:10.20944/preprints202605.1175.v1

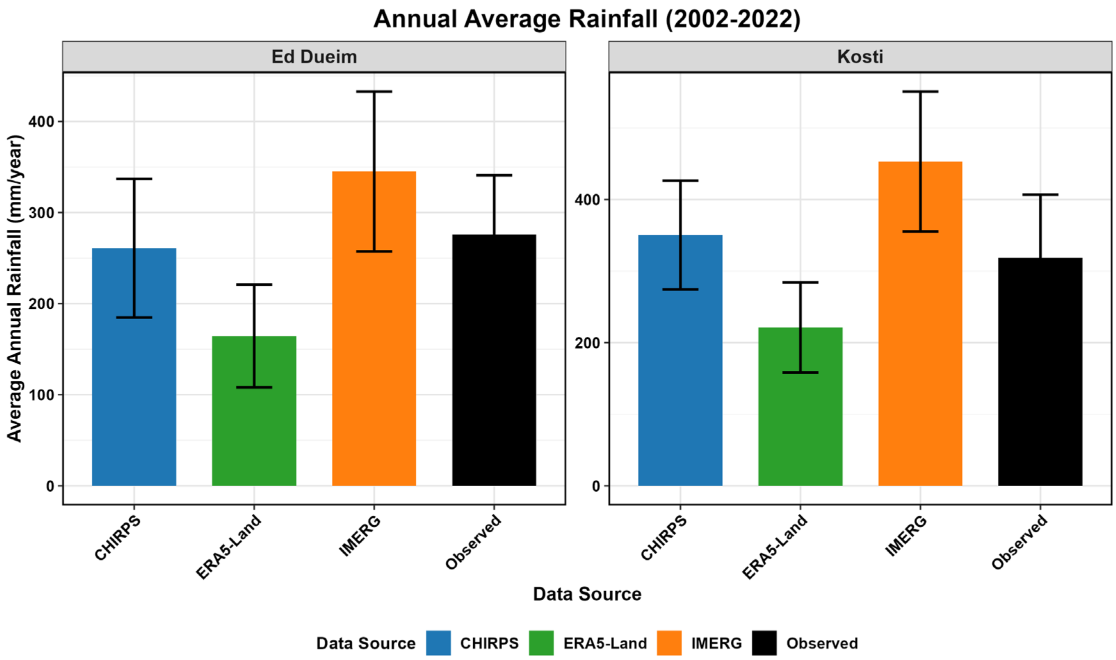

Figure 3. Mean annual rainfall (2002-2022) and associated interannual variability at Ed Dueim and Kosti for all products and gauge observations. Error bars denote one standard deviation of annual totals.

5.3. Statistical Agreement with In-Situ Observations

5.3.1. Kosti Station

When evaluated monthly, CHIRPS delivered its most reliable overall performance at Kosti (Table 3). With a Pearson correlation of 0.88 ($R^2 = 0.78$) and a Spearman's rho of 0.93, the data indicate that the product closely mirrors both the volume and the month-to-month timing of local rainfall. Error margins remain narrow, with an RMSE of 24.4 mm and an MAE of 11.9 mm. Furthermore, efficiency scores of 0.77 (NSE) and 0.79 (KGE) highlight how effectively CHIRPS captures the region's distinct seasonal pulse. A systematic negative bias of -12.8% is present but does not undermine the overall coherence of the monthly fit.

At the annual scale, CHIRPS skill declines markedly, as shown in (Figure 4). Pearson correlation drops to 0.39, and R^2 falls to 0.15; NSE turns negative (-0.13), indicating that CHIRPS marginally fails to outperform the observed long-term mean as a predictor of annual totals. This collapse in skill from monthly to annual scales is not incidental. The strong monthly performance is substantially sustained by the shared seasonal cycle, which all products capture well. Once this seasonal structure is effectively removed through annual aggregation, the products' ability to track genuine interannual variability is considerably weaker than monthly statistics alone would suggest.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1175.v1

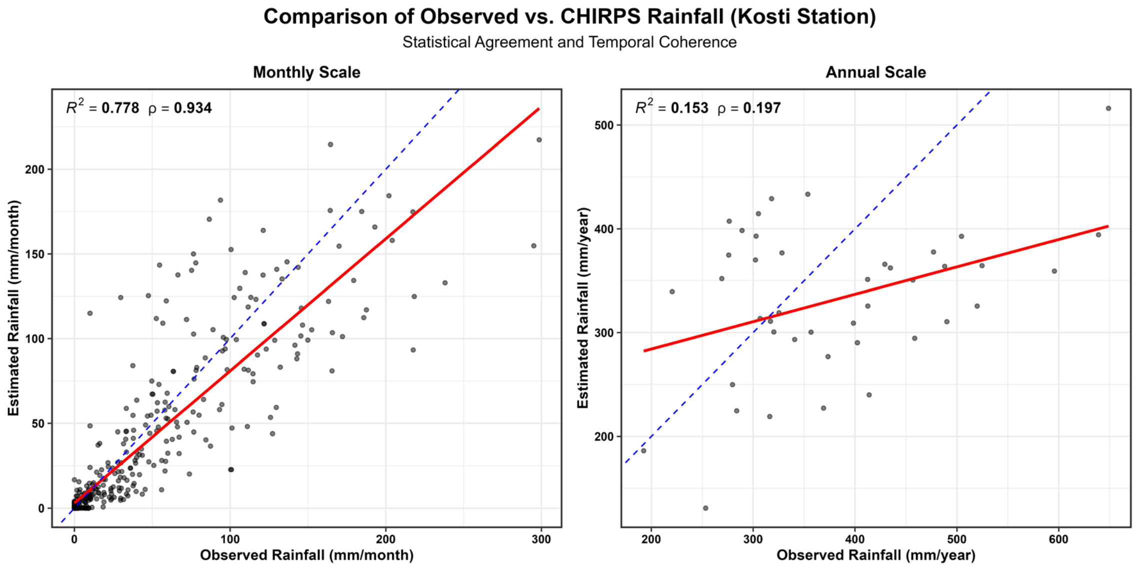

Figure 4. Statistical Agreement Between Observed and CHIRPS Rainfall at Kosti Station Across Monthly and Annual Timescales.

IMERG at Kosti achieves slightly higher monthly correlation statistics (Pearson r = 0.90, R^2 = 0.81, Spearman rho = 0.91) than CHIRPS, indicating strong temporal phase agreement. However, efficiency metrics diverge substantially; NSE = 0.53 and KGE = 0.44 fall well below CHIRPS values, reflecting the distorting effect of IMERG's +42.4% wet bias on absolute accuracy. At the annual scale, IMERG records the highest linear correlation (r = 0.66, R^2 = 0.43), indicating relative skill in distinguishing wet from dry years, yet its NSE drops to -2.21, confirming that its large systematic bias renders annual absolute totals far less reliable than simply using the long-term observed mean. Look at Figure 5.

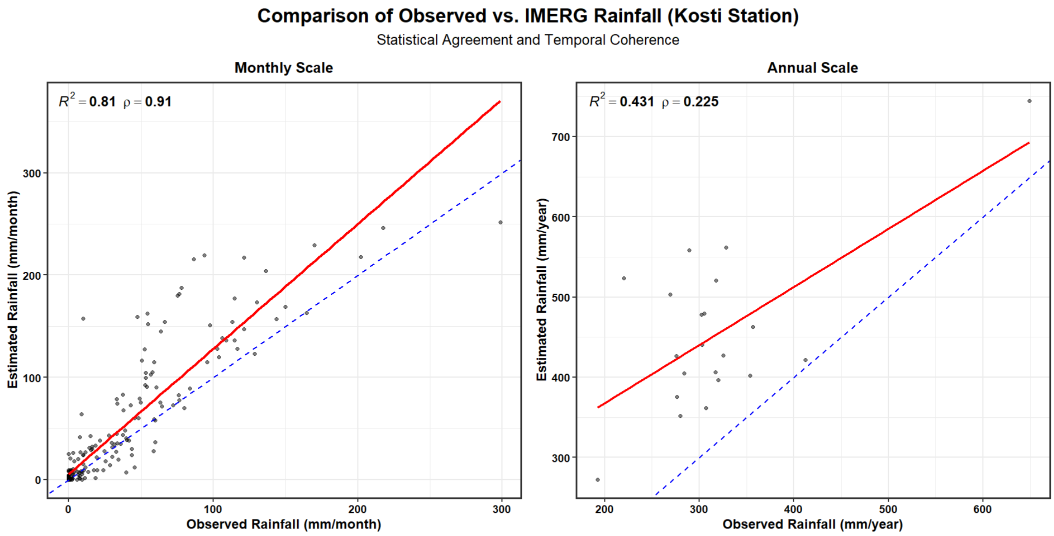

Figure 5. Statistical Agreement Between Observed and IMERG Rainfall at Kosti Station Across Monthly and Annual Timescales.

ERA5-Land lags behind both CHIRPS and IMERG at Kosti at the monthly scale. Its Pearson correlation of 0.56 and R^2 = 0.32 reflect only moderate linear agreement; as shown on (Figure 6) NSE = 0.16 sits marginally above zero, indicating near-negligible improvement over a climatological mean predictor. The RMSE of 51.2 mm is the highest of all products at Kosti, more than double that of CHIRPS. At the annual scale, performance deteriorates categorically; Pearson r = 0.29, R^2 = 0.09, NSE

= -7.14, and KGE = -1.02 are unambiguous indicators that ERA5-Land annual totals are far less reliable than the long-term mean by a wide margin. The annual RMSE of 314.4 mm (Table 3), more than twice that of IMERG and nearly three times that of CHIRPS, places ERA5-Land in a distinct performance category for annual analysis at this station.

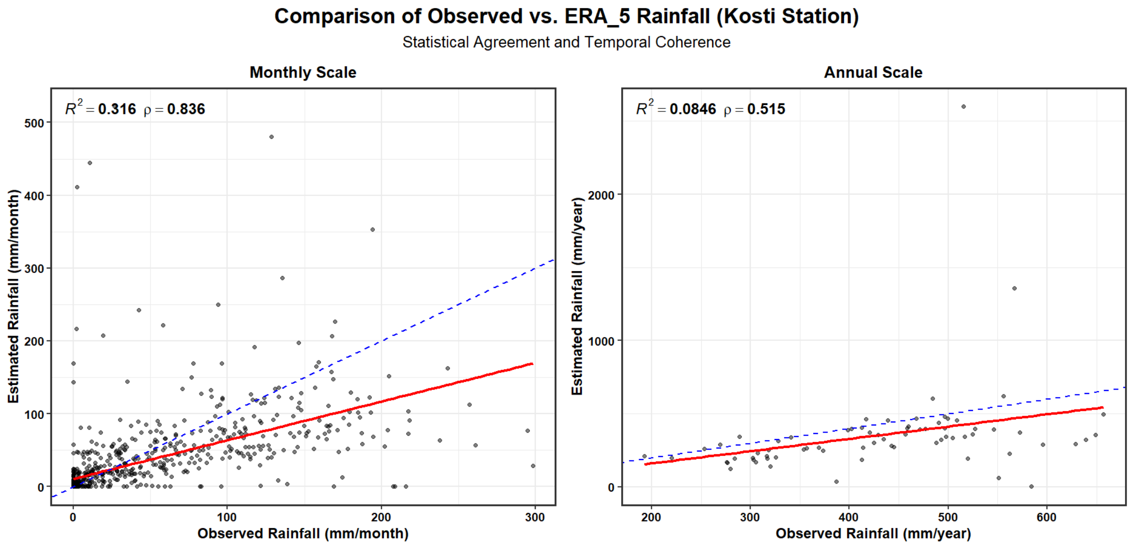

Figure 6. Statistical Agreement Between Observed and ERA-5 Rainfall at Kosti Station Across Monthly and Annual Timescales.

Table 3. Statistical performance metrics for CHIRPS, IMERG, and ERA5-Land against gauge observations at Kosti Station (monthly and annual timescales). RMSE and MAE in mm; rho = Spearman rank correlation.

Scale	Product	Period	rho	r	R^2	NSE	KGE	RMSE	MAE	Bias (mm)	Bias (%)
Monthly	CHIRPS	1981-2022	0.93	0.88	0.78	0.77	0.79	24.4	11.9	-2048.2	-12.8
	IMERG	2002-2022	0.91	0.90	0.81	0.53	0.44	29.8	14.4	+2834.1	+42.4
	ERA5-Land	1952-2022	0.84	0.56	0.32	0.16	0.52	51.2	22.5	-5406.6	-17.7
Annual	CHIRPS	1981-2022	0.20	0.39	0.15	-0.13	0.30	114.1	99.0	-2048.2	-12.8
	IMERG	2002-2022	0.23	0.66	0.43	-2.21	0.44	154.8	135.0	+2834.1	+42.4
	ERA5-Land	1952-2022	0.52	0.29	0.09	-7.14	-1.02	314.4	166.8	-5406.6	-17.7

5.3.2. Ed Dueim Station

The performance hierarchy at Ed Dueim mirrors that of Kosti, with all three products achieving higher skill at the monthly scale than at the annual scale (Table 4). CHIRPS demonstrates the strongest overall monthly agreement: Pearson r = 0.89, Spearman rho = 0.92, $R^2 = 0.79$, NSE = 0.77, and KGE = 0.88, with modest error metrics (RMSE = 19.0 mm; MAE = 8.2 mm) and a small systematic dry bias of -5.4%. This represents the best combined performance of any product at either station and the lowest bias of any product at Ed Dueim. At the annual scale, CHIRPS maintains the highest skill among all products (Pearson r = 0.59, $R^2 = 0.34$, NSE = 0.09); look at (Figure 7), the only product to achieve a non-negative annual NSE at either station, indicating marginally positive predictive skill at this timescale.

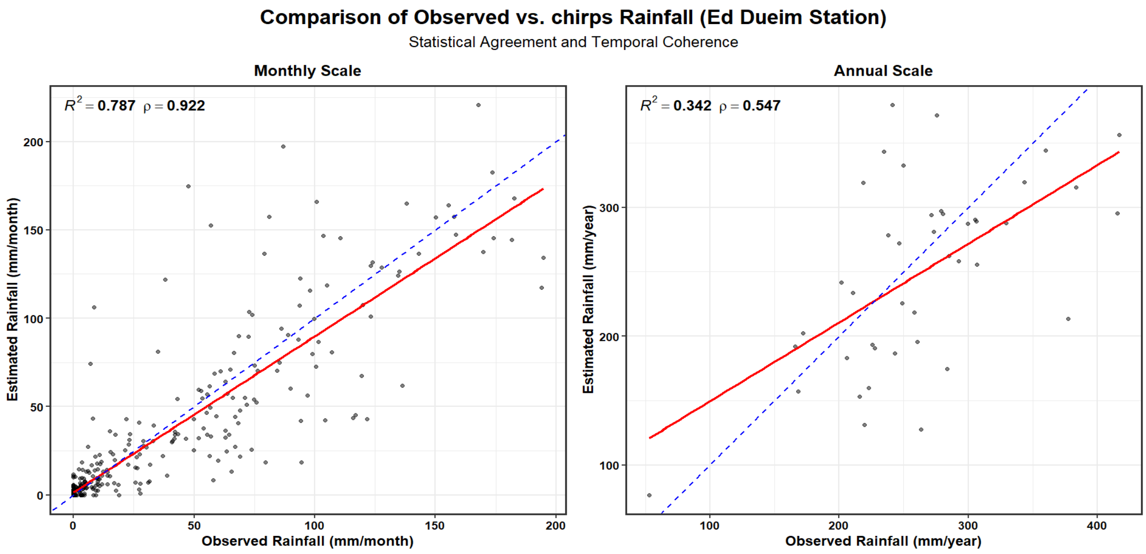

Figure 7. Statistical Agreement Between Observed and CHIRPS Rainfall at Ed Dueim Station Across Monthly and Annual Timescales.

IMERG achieves monthly correlations comparable to CHIRPS at Ed Dueim (Pearson r = 0.88, R^2 = 0.78, Spearman rho = 0.88), but NSE = 0.66 and KGE = 0.66 are noticeably lower, reflecting the effect of its +25.0% systematic wet bias on efficiency metrics. At the annual scale, IMERG's NSE of -1.93 confirms that absolute annual totals are substantially degraded relative to the long-term mean, a direct consequence of its persistent bias. Annual RMSE reaches 108.5 mm and MAE 86.0 mm, both well above CHIRPS equivalents (look at Figure 8 and Table 4).

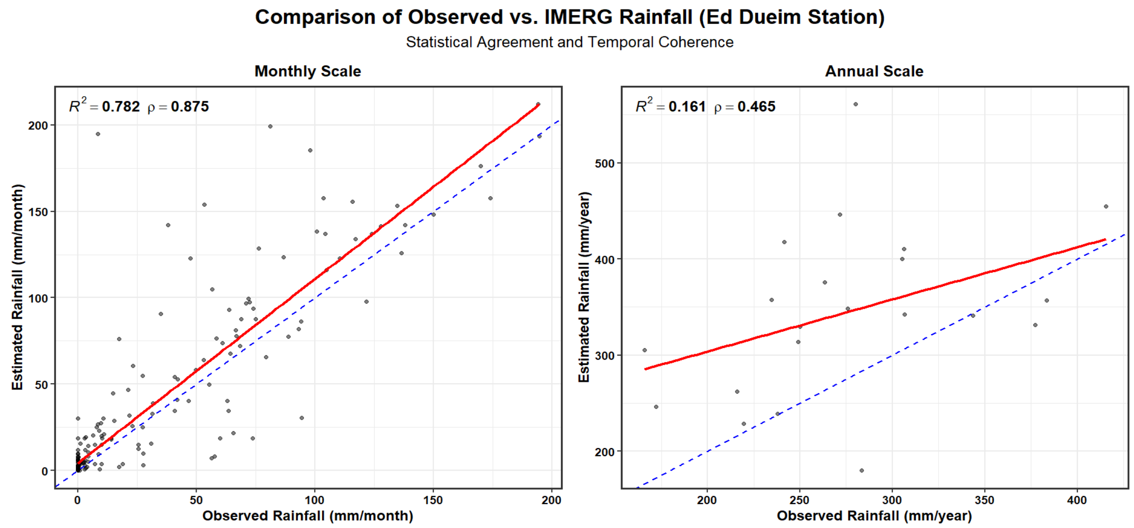

Figure 8. Statistical Agreement Between Observed and IMERG Rainfall at Ed Dueim Station Across Monthly and Annual Timescales.

ERA5-Land at Ed Dueim records a monthly Pearson correlation of only 0.44 and a negative NSE (-0.22), indicating that the reanalysis fails to outperform the observed monthly mean as a predictor even at the monthly scale. The Spearman rho of 0.79 reveals a revealing disconnect; ERA5-Land retains moderate capacity to rank wet and dry months correctly yet translates this ranking into actual rainfall amounts with poor fidelity. This divergence between rank skill and magnitude skill is a hallmark of the deep dry bias (PBIAS = -4.6%) and high scatter of the reanalysis at this station as shown on (Figure 8 and Table 4). At the annual scale, ERA5-Land's performance is effectively negligible Pearson r = 0.12, R^2 = 0.01, NSE = -15.34, KGE = -2.16. Annual RMSE climbs to 286.2 mm and MAE to 129.7 mm, yielding the weakest performance of any product at either station, as illustrated on (Figure 9 and Table 4).

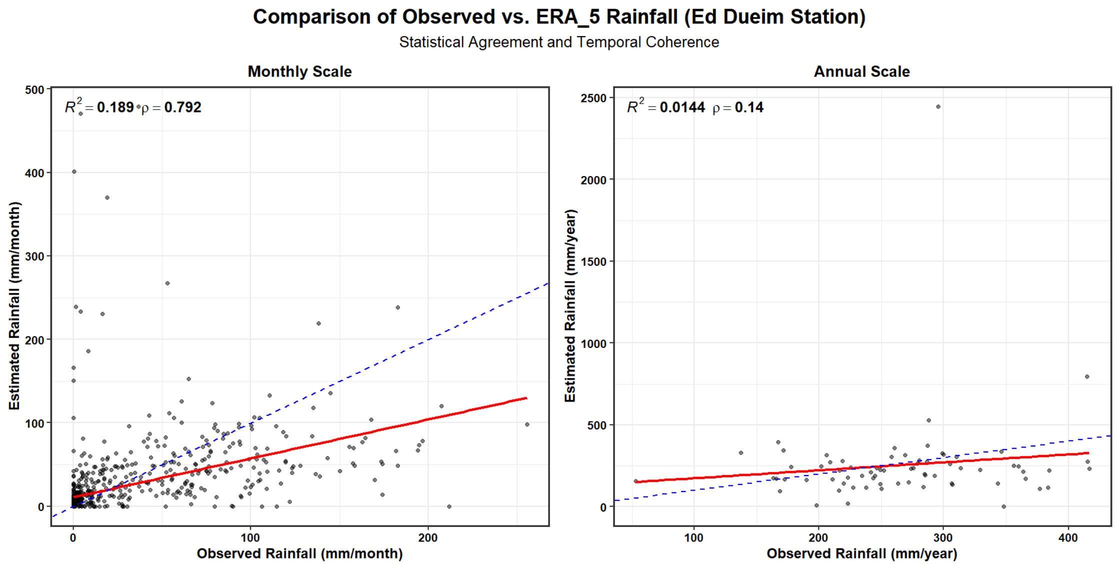

Figure 9. Statistical Agreement Between Observed and ERA-5 Land Rainfall at Ed Dueim Station Across Monthly and Annual Timescales.

Table 4. Statistical performance metrics for CHIRPS, IMERG, and ERA5-Land against gauge observations at Ed Dueim (monthly and annual timescales).

Scale	Product	Period	rho	r	R2	NSE	KGE	RMSE	MAE	Bias (mm)	Bias (%)
Monthly	CHIRPS	1981-2022	0.92	0.89	0.79	0.77	0.88	19.0	8.2	-595.8	-5.4
	IMERG	2002-2022	0.88	0.88	0.78	0.66	0.66	23.5	10.6	+1449.8	+25.0
	ERA5-Land	1952-2022	0.79	0.44	0.19	-0.22	0.43	45.2	17.9	-872.8	-4.6
Annual	CHIRPS	1981-2022	0.55	0.59	0.34	0.09	0.58	65.6	52.2	-595.8	-5.4
	IMERG	2002-2022	0.47	0.40	0.16	-1.93	0.26	108.5	86.0	+1449.8	+25.0
	ERA5-Land	1952-2022	0.14	0.12	0.01	-15.34	-2.16	286.2	129.7	-872.8	-4.6

5.4. Rainfall Detection Skill and Contingency Analysis

Beyond simply capturing rainfall volume, a gridded product's true utility in semi-arid hydrology often hinges on its ability to identify when rain actually occurs accurately. This is a critical baseline in regions where the monsoon is both brief and unpredictable (Ouatiki et al., 2024). We evaluated this detection skill across four distinct seasonal windows at both stations, with the detailed results compiled in (Table 5). To provide a clearer picture of these metrics, (Figure 10) visualizes categorical performance, while (Figure 11) employs a Roebber performance diagram to illustrate the integrated trade-off between the Probability of Detection (POD), the False Alarm Ratio (FAR), and the Critical Success Index (CSI).

5.4.1. Full-Period Performance

At the full-period scale, all three products achieve high POD at both stations (>= 0.88), confirming that the dominant monsoon signal is reliably detected regardless of product type. Differentiation emerges primarily through FAR. CHIRPS records the lowest full-period FAR at both stations (0.12 at Ed Dueim; 0.04 at Kosti), translating into the highest or joint-highest CSI values (0.79 and 0.84, respectively). ERA5-Land carries the highest full-period FAR at Ed Dueim (0.27), reflecting systematic reporting of rainfall events unconfirmed by the gauge. In the Roebber diagram (Figure

 doi:10.20944/preprints202605.1175.v1

10), CHIRPS and IMERG cluster near the high-POD, low-FAR region along favorable CSI contours (>= 0.8) during the rainy season, while ERA5-Land plots toward higher FAR values, providing visual confirmation of its greater false alarm burden.

5.4.2. Seasonal Detection: Winter, Pre-Monsoon, and Rainy Season

The most unambiguous finding in the contingency analysis concerns the winter period (November-February). Both ERA5-Land and IMERG record FAR = 1.00 and CSI = 0.00 at Ed Dueim during this season, generating multiple false alarms against zero confirmed hits. CHIRPS records zero events and zero false alarms in winter at both stations, the only product to correctly represent the dry season as rainfall-free. This categorical failure during the dry season has direct implications; any application relying on ERA5-Land or IMERG for dry-season water balance estimation, drought onset detection, or baseflow separation would be fundamentally compromised without prior correction.

During the rainy season (June-October), all products perform at their strongest. IMERG achieves the highest rainy-season POD at both stations (0.97), with CSI reaching 0.84 at Ed Dueim and 0.91 at Kosti. CHIRPS performs comparably (POD = 0.94, CSI = 0.84 at Ed Dueim; POD = 0.91, CSI = 0.89 at Kosti), and ERA5-Land achieves acceptable skill during this period (CSI = 0.75 and 0.86). FAR values are low across all products during the monsoon peak, most strikingly for CHIRPS at Kosti (FAR = 0.02).

The pre-monsoon period (March-May) reveals the sharpest performance contrast. At Ed Dueim, CHIRPS records a pre-monsoon POD of only 0.09 and a CSI of 0.09, effectively missing almost all early-season events. ERA5-Land achieves a higher POD (0.62) but at the cost of severe overreporting, with FAR = 0.80 and a frequency bias of 3.15, reporting more than three times as many rain events as the gauge records. IMERG shows moderate pre-monsoon skill (POD = 0.44, FAR = 0.43, CSI = 0.33 at Ed Dueim). Performance improves at the wetter Kosti station across all products, with CHIRPS reaching POD = 0.68 and CSI = 0.58, suggesting that pre-monsoon detection is more reliable where events are more frequent and spatially coherent.

The elevated FAR values during the pre-monsoon period, particularly for ERA5-Land, are consistent with several plausible mechanisms in this environment, including sub-grid convective parameterization errors generating column precipitation that does not reach the surface; the well-documented virga effect, whereby precipitation evaporates within the deep dry sub-cloud layer characteristic of pre-monsoon central Sudan (Knippertz et al., 2009); and the fundamental scale mismatch between point gauge observations and gridded areal estimates during spatially localized convective events (Hossain & Anagnostou, 2004).

doi:10.20944/preprints202605.1175.v1

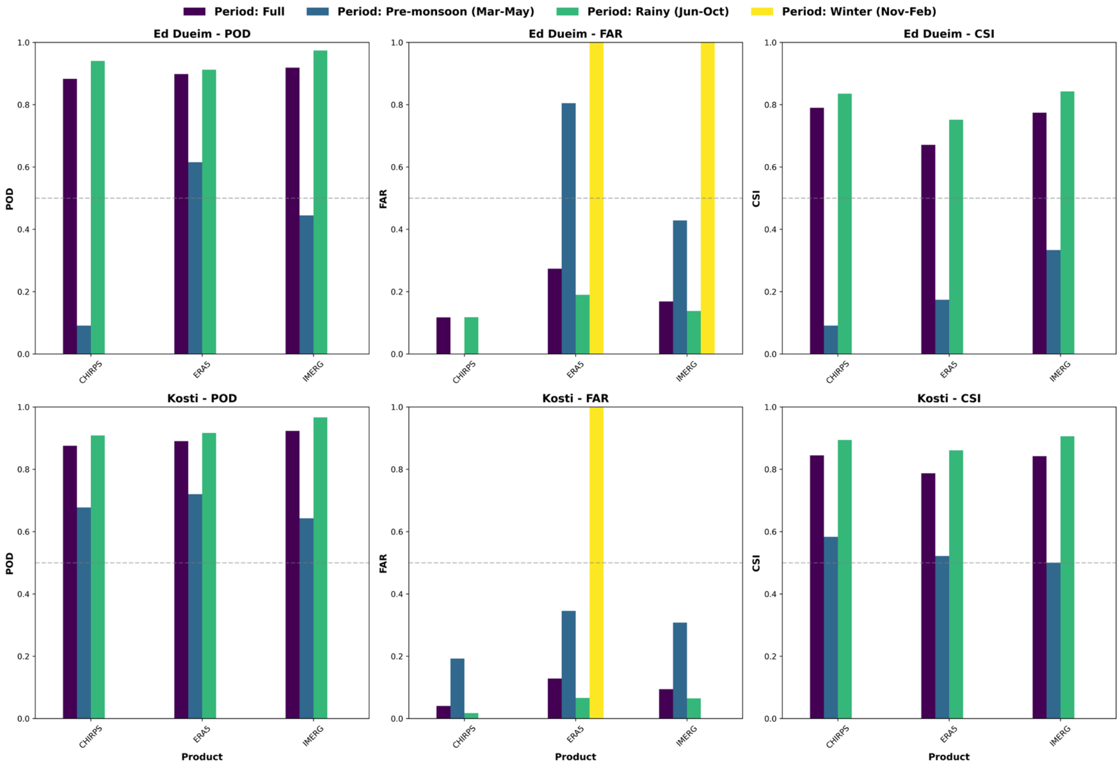

Figure 10. Seasonal rainfall detection skill scores (POD, FAR, CSI) for CHIRPS, ERA5-Land, and IMERG at Ed Dueim and Kosti across seasonal periods.

Table 5. Rainfall event detection performance metrics (POD, FAR, CSI, frequency bias) for CHIRPS, ERA5-Land, and IMERG at Ed Dueim and Kosti, stratified by season.

	Product	Period	POD	FAR	CSI	BIAS	Hits	False Alarms
El Dueim Station	CHIRPS	Full	0.88	0.12	0.79	1.00	143	19
		Winter (Nov-Feb)					0	0
		Pre-monsoon (Mar-May)	0.09	0.00	0.09	0.09	1	0
		Rainy (Jun-Oct)	0.94	0.12	0.84	1.07	142	19
	ERA5	Full	0.90	0.27	0.67	1.24	247	93
		Winter (Nov-Feb)		1.00	0.00		0	4
		Pre-monsoon (Mar-May)	0.62	0.80	0.17	3.15	8	33
		Rainy (Jun-Oct)	0.91	0.19	0.75	1.13	239	56
	IMERG	Full	0.92	0.17	0.77	1.10	79	16
		Winter (Nov-Feb)		1.00	0.00		0	1
		Pre-monsoon (Mar-May)	0.44	0.43	0.33	0.78	4	3
		Rainy (Jun-Oct)	0.97	0.14	0.84	1.13	75	12
Kosti Station	CHIRPS	Full	0.88	0.04	0.84	0.91	190	8
		Winter (Nov-Feb)					0	0
		Pre-monsoon (Mar-May)	0.68	0.19	0.58	0.84	21	5
		Rainy (Jun-Oct)	0.91	0.02	0.89	0.92	169	3
	ERA5	Full	0.89	0.13	0.79	1.02	333	49
		Winter (Nov-Feb)		1.00	0.00		0	9
		Pre-monsoon (Mar-May)	0.72	0.35	0.52	1.10	36	19
		Rainy (Jun-Oct)	0.92	0.07	0.86	0.98	297	21
	IMERG	Full	0.92	0.09	0.84	1.02	96	10
		Winter (Nov-Feb)					0	0

Pre-monsoon (Mar-May)	0.64	0.31	0.50	0.93	9	4
Rainy (Jun-Oct)	0.97	0.06	0.91	1.03	87	6

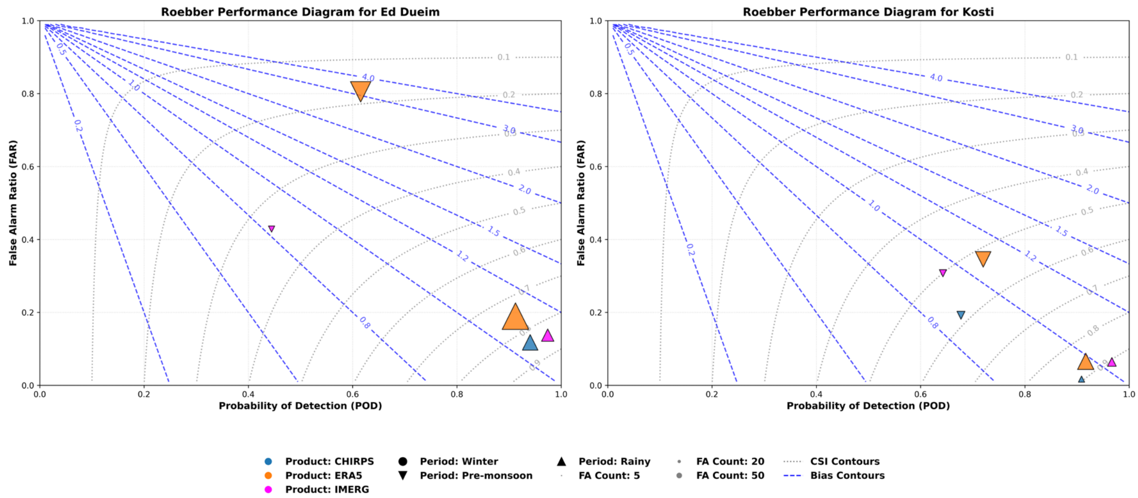

Figure 11. Roebber Performance Diagrams for rainfall event detection at Ed Dueim and Kosti, displaying the integrated POD-FAR-CSI trade-off for each product across seasonal periods.

5.5. Integrated Performance Synthesis

The synthesizes monthly-scale performance across KGE, NSE, R2, and Spearman rho at both stations, providing the clearest integrated visual summary of the multi-metric evidence. The product hierarchy is consistent across both stations and across all four metrics. CHIRPS leads at Ed Dueim (monthly KGE = 0.88, NSE = 0.77, R^2 = 0.79, Spearman rho = 0.92) and maintains this position at Kosti (KGE = 0.79, NSE = 0.77, R^2 = 0.78, Spearman rho = 0.83). IMERG occupies an intermediate position, with acceptable monthly correlation statistics but substantially degraded efficiency metrics driven by its persistent wet bias. ERA5-Land records the weakest profile at both stations across all four metrics, with monthly NSE at or below zero, as illustrated on (Figure 12).

This ranking holds when extended to the full evidence base. CHIRPS produces the smallest and most consistent mean annual bias (-5.5% at Ed Dueim, +10.1% at Kosti), compared to IMERG's systematic overestimation (+25.0% and +42.4%) and ERA5-Land's persistent underestimation (-40.4% and -30.5%). On detection skill, CHIRPS achieves the lowest FAR at both stations and is the only product to correctly represent the dry season as rainfall-free. At the annual scale, no product demonstrates meaningful predictive skill at either station; all NSE values are negative or negligible, but this is a shared structural limitation of all three products rather than a differentiating characteristic.

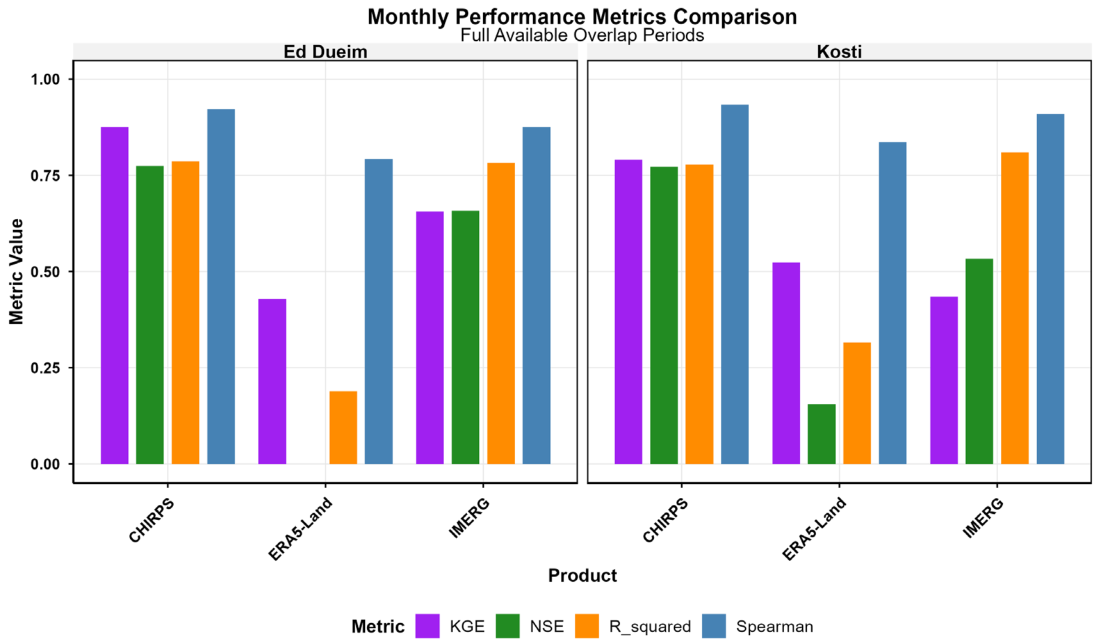

Figure 12. Integrated monthly-scale performance plots for CHIRPS, IMERG, and ERA5-Land at Ed Dueim and Kosti, showing KGE, NSE, R2, and Spearman rho.

6. Discussion

6.1. Comparative Performance and Mechanistic Interpretation

The systematic performance advantage of CHIRPS over IMERG and ERA5-Land at both stations across all temporal scales and metric types is consistent with the findings of prior Satellite Rainfall Estimation (SREs) evaluation studies in the Sahel and sub-Saharan Africa (Toté et al., 2015; Dinku et al., 2018; Satgé et al., 2020). The primary mechanistic advantage of CHIRPS in this environment is its gauge-merging algorithm; by incorporating station observations directly into the product construction, CHIRPS anchors its estimates to the observed climatological signal in a manner that reduces systematic bias at gauge locations. This partially explains why the CHIRPS bias, while not negligible, is more stable in sign, smaller in magnitude, and more consistent between stations than the biases of IMERG and ERA5-Land.

The persistent wet bias of IMERG across the region is a well-documented characteristic of passive microwave retrieval algorithms operating over semi-arid surfaces with high land surface emissivity. In these environments, the microwave emission from warm, dry soils can be misinterpreted by retrieval algorithms as a precipitation signal, contributing to false rainfall detection and magnitude overestimation (Maggioni et al., 2016). The strong rank correlation of IMERG at both stations, particularly at the annual scale, where it achieves r = 0.66 at Kosti, is therefore not inconsistent with a large systematic bias, the product correctly identifies the relative ordering of wet and dry periods, but its absolute estimates are consistently and substantially inflated.

ERA5-Land's underperformance relative to SREs in this region reflects fundamental limitations of the convective parameterization schemes employed in reanalysis systems in semi-arid convective environments. Convective rainfall events in the Sudan Sahel are characterized by highly localized mesoscale convective systems that operate at spatial scales finer than the ERA5-Land grid, rendering their accurate simulation inherently challenging (Vischel et al., 2019). The categorical winter-period failure of ERA5-Land generating multiple false alarms in a season with zero observed gauge events further reflects a model-level tendency to produce spurious precipitation in extremely dry atmospheric states, a bias pattern identified in other ECMWF reanalysis generations across the Sahel (Nikulin et al., 2012).

 doi:10.20944/preprints202605.1175.v1

6.2. Temporal Scale-Dependence of Product Skill

The systematic degradation of statistical skill from monthly to annual timescales - observed across all three products at both stations - is a critical finding with direct methodological implications. At the monthly scale, strong correlations (CHIRPS r ~ 0.88-0.89; IMERG r ~ 0.88-0.90) are in large part a consequence of the shared seasonal cycle; all products and the gauge record exhibit low values from November to May and high values from June to October. This shared structure inflates correlation statistics without necessarily implying accurate reproduction of sub-seasonal variability. Once annual aggregation removes this seasonal signal, the products' true capacity to capture genuine interannual variability is revealed - and at both stations, this capacity is weak to negligible across all products evaluated.

This temporal scale-dependence has been reported in analogous semi-arid settings ((Teutschbein & Seibert, 2012; Ouatiki et al., 2024) and represents a fundamental constraint on the applicability of multi-year drought and trend assessments derived from these products without prior bias correction. The finding that even CHIRPS - the best-performing product at the monthly scale - achieves an annual NSE of only +0.09 at Ed Dueim and -0.13 at Kosti in its raw, uncorrected form underscores that monthly validation alone is insufficient for justifying product use in annual or decadal applications.

6.3. Implications for Operational Product Selection and Bias Correction

The operational implications of this evaluation are directly structured by the consistency and predictability of each product's bias profile. CHIRPS exhibits a systematic dry bias of 5-13% that is stable in sign across both stations and consistent across seasons; this is precisely the error profile that quantile mapping and linear scaling bias correction approaches are most effective at reducing (Teutschbein & Seibert, 2012). Bias-correction procedures applied to products with stable, well-characterized offsets are substantially more transferable and robust than those targeting products whose bias direction or magnitude shifts between locations or seasons.

IMERG retains operational relevance in a specific and bounded context, near-real-time detection of rainfall occurrence during the active monsoon season, where its rainy-season POD is the highest of all products (0.97 at both stations) and its monthly correlation is comparable to CHIRPS'. This characteristic high detection sensitivity with systematic magnitude overestimation renders IMERG better suited to threshold-based early warning applications (flood onset, crop water stress detection) than to quantitative water balance estimation. Particularly for Monsoon Window and particularly for annual, drought, or dry-season applications, IMERG's large mean bias and categorical winter failures render it unsuitable without rigorous correction.

ERA5-Land, in its raw, uncorrected form, is not recommended for precipitation-based applications at either station in White Nile State. Its principal operational value in this region lies in its temporal depth extending back to 1952, which, combined with rigorous bias correction anchored to the gauge record, may support historical reconstruction of rainfall anomalies for periods preceding the satellite era. This application context aligns with the established use of reanalysis products as physically consistent dynamical "proxies" for pre-instrumental periods rather than as direct precipitation estimates (Dee et al., 2011).

6.4. Limitations and Future Directions

This evaluation is constrained by the availability of only two long-term gauge stations within White Nile State, limiting the spatial representativeness of conclusions. Point-based gauge observations inevitably reflect micro-scale variability that may not be representative of the areal rainfall captured by gridded products, a fundamental scale mismatch that contributes to apparent product error, particularly during spatially localized convective events (Hossain and Anagnostou, 2004). The absence of dense gauge networks or high-resolution remote sensing (Disdrometer,

 doi:10.20944/preprints202605.1175.v1

weather radar) in this region precludes a rigorous separation of scale mismatch from genuine product errors.

The analyses presented here are limited to raw, uncorrected product estimates. Future work should systematically evaluate the performance of bias-corrected CHIRPS and IMERG products at multiple spatial scales relevant to hydrological modeling in White Nile State and assess whether the temporal scale dependence of product skill is reduced following correction. The extension of this evaluation framework to temperature products, particularly CHIRTS for extreme temperature monitoring, represents an important complementary step for comprehensive hydro-climatic monitoring capacity in the region.

7. Conclusions

This study has conducted the most comprehensive, multi-metric evaluation of satellite precipitation products and reanalysis data for White Nile State, central Sudan, currently available in the literature. The following principal conclusions are drawn from the analysis:

1. All three gridded products - CHIRPS, IMERG, and ERA5-Land - accurately reproduce the unimodal June-October monsoon structure and the maximum August rainfall at both Ed Dueim and Kosti stations. Agreement in seasonal timing is a necessary but insufficient criterion for product suitability; systematic magnitude biases differentiate products substantially and have direct consequences for quantitative hydrological applications.

2. CHIRPS is the most suitable product for operational precipitation monitoring in White Nile State. Its monthly NSE (0.77 at both stations), KGE (0.79-0.88), and Spearman correlation (0.92-0.93) represent the strongest performance of any product across the multi-metric framework, and its stable 5-13% dry bias is structurally amenable to operational bias correction.

3. IMERG demonstrates high monthly correlation and the strongest rainy-season detection sensitivity (POD = 0.97), but overestimates mean annual rainfall by 25-42% and produces categorical dry-season failures (FAR = 1.00 in winter). Its use should be restricted to near-real-time monsoon season event detection, and bias correction is mandatory before any quantitative water balance application.

4. ERA5-Land, in its raw form, is not suitable for precipitation-based applications at either station. Monthly NSE values at or below zero, annual NSE reaching -15.34, systematic dry-season false alarms, and the weakest performance across all categorical metrics across all seasons collectively disqualify it from operational use without rigorous bias correction. Its primary value is as a temporally deep pronxy for pre-satellite climate reconstruction when corrected against the gauge record.

5. All products exhibit systematic skill degradation from monthly to annual timescales. No product achieves meaningful positive NSE at the annual scale, confirming that monthly-scale validation statistics alone are insufficient to justify product use in annual water balance, drought, or long-term trend applications. Bias correction targeted to stabilize interannual performance is an essential prerequisite for such applications.

These findings provide an evidence-based foundation for product selection, bias correction framework design, and the strategic integration of multiple data sources in hydroclimatic monitoring and water resource management across White Nile State and analogous semi-arid Sahelian environments.

Author Contributions: Conceptualization, A.Y.F.A.; methodology, A.Y.F.A.; software, A.Y.F.A.; validation, A.Y.F.A.; Z.G and. P.K; formal analysis, A.Y.F.A.; writing original draft preparation, A.Y.F.A.; writing review and editing, A.Y.F.A., Z.G., and P.K.; visualization, A.Y.F.A.; supervision, Z.G. and P.K. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: CHIRPS v2.0 data are publicly available from the Climate Hazards Center (https://www.chc.ucsb.edu/data/chirps). IMERG Final Run data are available from the NASA Goddard Earth Sciences Data and Information Services Center (https://disc.gsfc.nasa.gov). ERA5-Land data are available from

 doi:10.20944/preprints202605.1175.v1

the Copernicus Climate Change Service (https://cds.climate.copernicus.eu). In-situ gauge data from White Nile State were obtained from the Sudan Meteorological Authority and are available on reasonable request subject to data sharing agreements.

Ethics approval and consent to participate: Not applicable.

Acknowledgments: The authors gratefully acknowledge the support of the Stipendium Hungaricum Scholarship and the University of Sopron. Data was provided by the Climate Hazards Center (CHIRPS), NASA (IMERG), and the Copernicus Climate Change Service (ERA5-Land). In-situ gauge data were obtained from the Sudan Meteorological Authority and are available upon reasonable request.

Conflicts of Interest: the authors declare no conflicts of interest.

References

Adam, A. Y. F., Gribovszki, Z., & Kalicz, P. (2025). Enhancing Rainfall Measurement Using Remote Sensing Data in Sparse Rain Gauge Networks: A Case Study in White Nile State, Sudan †. *Engineering Proceedings*, *94*(1). https://doi.org/10.3390/ENGPROC2025094019

Bell, M. A., Lamb, P. J., Bell, M. A., & Lamb, P. J. (2006). Integration of Weather System Variability to Multidecadal Regional Climate Change: The West African Sudan–Sahel Zone, 1951–98. *Journal of Climate*, *19*(20), 5343–5365. https://doi.org/10.1175/JCLI4020.1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., … Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/QJ.828;REQUESTEDJOURNAL:JOURNAL:1477870X;PAGE:STRING:ARTICLE/CHAPTER

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018a). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*, *144*, 292–312. https://doi.org/10.1002/QJ.3244;PAGEGROUP:STRING:PUBLICATION

Elagib, N. A. (2010). Trends in intra-and inter-annual temperature variabilities across Sudan. *Ambio*, *39*(5), 413–429. https://doi.org/10.1007/s13280-010-0042-3

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066-. https://doi.org/10.1038/SDATA.2015.66;SUBJMETA

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/J.JHYDROL.2009.08.003

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*, *4*(2), 135–143. https://doi.org/10.1061/(asce)1084-0699(1999)4:2(135)

Hossain, F., & Anagnostou, E. N. (2004). Assessment of current passive-microwave- and infrared-based satellite rainfall remote sensing for flood prediction. *Journal of Geophysical Research: Atmospheres*, *109*(7). https://doi.org/10.1029/2003JD003986

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K. L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B., & Xie, P. (2020a). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). *Advances in Global Change Research*, *67*, 343–353. https://doi.org/10.1007/978-3-030-24568-9_19/FIGURES/3

Jolliffe, I. T., & Stephenson, D. B. (2012). Forecast Verification: A Practitioner's Guide in Atmospheric Science: SECOND EDITION. Forecast Verification: A Practitioner's Guide in Atmospheric Science, Second Edition, 1–274. https://doi.org/10.1002/9781119960003

Knippertz, P., Trentmann, J., & Seifert, A. (2009). High-resolution simulations of convective cold pools over the northwestern Sahara. *Journal of Geophysical Research Atmospheres*, *114*(8). https://doi.org/10.1029/2008JD011271

Maggioni, V., Sapiano, M. R. P., Adler, R. F., Maggioni, V., Sapiano, M. R. P., & Adler, R. F. (2016). Estimating Uncertainties in High-Resolution Satellite Precipitation Products: Systematic or Random Error? *Journal of Hydrometeorology*, *17*(4), 1119–1129. https://doi.org/10.1175/JHM-D-15-0094.1

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J. N. (2021a). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/ESSD-13-4349-2021

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J., Hänsler, A., Meijgaard, E. van, Samuelsson, P., Sylla, M. B., Sushama, L., Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., … Sushama, L. (2012). Precipitation Climatology in an Ensemble of CORDEX-Africa Regional Climate Simulations. *Journal of Climate*, *25*(18), 6057–6078. https://doi.org/10.1175/JCLI-D-11-00375.1

Ouatiki, H., Shit, T., Chehbouni, A., Ouatiki, H., Shit, T., Chehbouni, A., Ouatiki, Hamza, Shit, Taoufik, Chehbouni, & Abdelghani. (2024). Enhancing the rain/no-rain identification in remote sensing-based rainfall products: what impact on streamflow simulation. *EGUGA*, 20083. https://doi.org/10.5194/EGUSPHERE-EGU24-20083

Rauch, M., Bliefernicht, J., Maranan, M., Fink, A. H., & Kunstmann, H. (2025). Interannual Rainfall Variability in West Africa: Reconstruction Based on Atmospheric Circulation Patterns. *International Journal of Climatology*, *45*(14), e70088. https://doi.org/10.1002/JOC.70088;REQUESTEDJOURNAL:JOURNAL:10970088;JOURNAL:JOURNAL:10970088A;ISSUE:ISSUE:DOI

Salih, A. A. M., Zhang, Q., & Tjernström, M. (2015). Lagrangian tracing of Sahelian Sudan moisture sources. *Journal of Geophysical Research*, *120*(14), 6793–6808. https://doi.org/10.1002/2015JD023238

Salih, W., Epule, T. E., EL Khalki, E. M., Ouatiki, H., Erraki, S., Achli, S., & Chehbouni, A. (2023). A comprehensive assessment of satellite precipitation products over a semi-arid region: focus on extreme events. *Natural Hazards 2023 120:3*, *120*(3), 3037–3065. https://doi.org/10.1007/S11069-023-06317-Y

Satgé, F., Hussain, Y., Molina-Carpio, J., Pillco, R., Laugner, C., Akhter, G., & Bonnet, M. P. (2020). Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. *International Journal of Climatology*, *41*(S1), E517–E536. https://doi.org/10.1002/JOC.6704

Spearman, C. (1961). The Proof and Measurement of Association Between Two Things. *Studies in Individual Differences: The Search for Intelligence.*, 45–58. https://doi.org/10.1037/11491-005.

Teutschbein, C., & Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, *456–457*, 12–29. https://doi.org/10.1016/J.JHYDROL.2012.05.052

Toté, C., Patricio, D., Boogaard, H., van der Wijngaart, R., Tarnavsky, E., & Funk, C. (2015). Evaluation of Satellite Rainfall Estimates for Drought and Flood Monitoring in Mozambique. *Remote Sensing 2015, Vol. 7, Pages 1758-1776*, *7*(2), 1758–1776. https://doi.org/10.3390/RS70201758

Vischel, T., Panthou, G., Peyrillé, P., Roehrig, R., Quantin, G., Lebel, T., Wilcox, C., Beucher, F., & Budiarti, M. (2019). Precipitation Extremes in the West African Sahel: Recent Evolution and Physical Mechanisms. *Tropical Extremes: Natural Variability and Trends*, 95–138. https://doi.org/10.1016/B978-0-12-809248-4.00004-2

Willmott, C. J., Ackleson, S. G., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., O'Donnell, J., & Rowe, C. M. (1985). Statistics for the evaluation and comparison of models. *Journal of Geophysical Research: Oceans*, *90*(C5), 8995–9005. https://doi.org/10.1029/jc090ic05p08995

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

 doi:10.20944/preprints202605.1175.v1

Wilks, D. S. (2011). Statistical Methods in the Atmospheric Sciences. 697

Woods, D., Kirstetter, P. E., Vergara, H., Duarte, J. A., & Basara, J. (2023). Hydrologic Evaluation of the Global Precipitation Measurement Mission over the U.S.: Error Budget Analysis. *Journal of Hydrology*, *626*, 130212–130212. https://doi.org/10.1016/J.JHYDROL.2023.130212