European Journal of Agronomy 182 (2027) 128344

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Propagation of rainfall dataset differences into crop model outputs and economic indicators

Jacob Emanuel Joseph [a,c,*], Christoph Gornott [d,e], Rahel Laudien [d], Stefan Siebert [f], Omonlola Nadine Worou [b], Anthony M. Whitbread [c], Reimund P. Rötter [a,g]

[a] University of Göttingen, Tropical Plant Production and Agricultural Systems Modelling (TROPAGS), Göttingen, Germany
[b] International Livestock Research Institute (ILRI), Dakar, Senegal
[c] International Livestock Research Institute (ILRI), Dar es salaam, Tanzania
[d] Potsdam Institute for Climate Impact Research (PIK), Leibniz Association, Potsdam, Germany
[e] University of Kassel, Agroecosystem Analysis and Modelling, Witzenhausen, Germany
[f] University of Göttingen, Agronomy Group, Göttingen, Germany
[g] University of Göttingen, Campus Center of Biodiversity and Sustainable Land Use (CBL), Göttingen, Germany

ARTICLE INFO

Keywords:
Agricultural Production Systems sIMulator (APSIM)
Climate variability
Gridded precipitation products (GPPs)
Rainfall variability
Semi-arid maize systems
Yield response
Sensitivity analysis

ABSTRACT

Gridded precipitation products (GPPs) are widely used in agricultural analyses in data-scarce regions, yet their suitability is often assessed using meteorological performance alone. This study evaluates how differences among six GPPs (CHIRPS, CPC, ERA5, MERRA2, MSWEP, and TAMSAT) propagate into simulated maize yield and economic indicators, using the Agricultural Production Systems sIMulator (APSIM) in two contrasting semi-arid systems, Kaffrine (Senegal) and Kongwa (Tanzania). By linking GPP evaluation to crop-model propagation and economic analysis within a single framework, this study offers a more comprehensive basis for assessing GPP suitability than meteorological validation alone.

Rainfall performance varied across indicators and sites, with higher agreement for seasonal totals than onset, and higher detection skill for dry days than moderate or heavy rainfall. Several GPPs over-represented light rainfall and under-represented heavy rainfall, reflecting spatial averaging. These differences translated into site-specific yield responses driven by early-season rainfall: in Kaffrine, yield and nitrogen responses remained similar across datasets (KGE = 0.05–0.35), whereas in Kongwa, yield variability and optimal input levels were more sensitive to dataset choice (KGE = −0.24–0.60), particularly under wetter conditions.

Across sites, yield response shapes to nitrogen and planting density remained similar, but economically optimal input levels varied across datasets, indicating that agreement in rainfall indicators alone is insufficient to infer stability of crop model outputs. Evaluating GPP suitability for agronomic applications requires crop model diagnostics alongside meteorological validation, with explicit environmental context, as demonstrated here within a single process-based model (APSIM) across two semi-arid maize systems.

1. Introduction

Agriculture in semi-arid regions is increasingly affected/threatened by climate variability and more frequent adverse weather events, with direct impacts on crop yields and food security (Alimagham et al., 2024; Frieler et al., 2017). The semi-arid zones of Senegal and Tanzania are particularly vulnerable due to limited and erratic rainfall (Tonnang et al., 2017). Climate-related risks include shifts in seasonal rainfall patterns, such as delayed rainy seasons, as well as an increased frequency of extreme events, including prolonged droughts and heavy precipitation, all of which adversely affect crop yields. (Gwambene et al., 2022; MacCarthy et al., 2021). These climatic pressures are compounded by biotic stresses such as pests and diseases (Deutsch et al., 2018; Tonnang et al., 2017), further straining smallholder production systems.

Accurate rainfall information is fundamental for agricultural planning and risk management in rainfed systems (Ogunbode et al., 2022). Yet many regions rely on sparse and unevenly distributed weather stations that fail to capture the spatial and temporal variability of rainfall needed for localized decision-making (Dinku, 2019; Stern and Cooper,

* Corresponding author at: University of Göttingen, Tropical Plant Production and Agricultural Systems Modelling (TROPAGS), Göttingen, Germany.
E-mail address: J.Emanuel@cgiar.org (J.E. Joseph).

https://doi.org/10.1016/j.eja.2026.128344
Received 10 April 2026; Received in revised form 22 July 2026; Accepted 17 September 2026
Available online 24 September 2026
1161-0301/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Open access funding enabled and organized by Projekt DEAL.

2011). This limits the ability to estimate rainfall onset and cessation, identify dry spells, and guide management decisions related to planting dates, fertilizer application, and planting densities. High-resolution gridded precipitation products (GPPs) are designed to address these gaps, integrating satellite observations, ground measurements, and modelling techniques to provide spatially continuous rainfall datasets (Sun et al., 2018). GPPs are increasingly used in climate risk assessment and agricultural planning, offering improved spatiotemporal coverage in data-scarce regions (Kessabi et al., 2023).

However, the performance of GPPs varies across regions due to differences in climate, topography, and input data sources (Chatterjee et al., 2023; Nassaj et al., 2022). Studies highlight substantial disparities in their ability to capture rainfall indicators such as seasonal totals, rainfall intensity, and onset timing (Zandler et al., 2019; Ahmed et al., 2024), underscoring the need for region-specific validation.

While many evaluations focus on rainfall indicators alone, few extend to crop modelling or agricultural impact assessment. Most studies also stop at simulated yields and do not examine whether rainfall dataset choice alters agronomic recommendations or economic outcomes. Gummadi et al. (2022), for instance, evaluated five satellite rainfall products against a DSSAT maize model in Vietnam but assessed only yield-level agreement (RMSE, correlation), without testing management or economic implications. Josephson et al. (2026) link rainfall product choice to economic outcomes, but through reduced-form statistical relationships rather than a process-based crop model, and do not test whether rainfall differences alter agronomic decisions such as optimal nitrogen rate or planting density. To our knowledge, no study has combined process-based crop model propagation with a factorial agronomic sensitivity analysis and economic translation across contrasting rainfall regimes. This gap constrains confidence in using GPP-driven crop–climate simulations for crop-model-based analysis in smallholder systems. Addressing this requires systematic diagnostics that link rainfall dataset differences to crop model processes and their resulting responses under historical climate variability.

Building on the need to assess the reliability of climate data for agricultural analyses, this study evaluates the robustness of crop model simulations to different rainfall datasets in semi-arid smallholder systems. Rather than assessing GPPs solely on rainfall indicators, the analysis links rainfall diagnostics with process-based crop simulations to determine how differences in the rainfall dataset alter simulated crop performance and derived economic indicators. The Agricultural Production Systems sIMulator (APSIM) was selected over alternative process-based crop models, such as DSSAT and AquaCrop, for its explicit representation of daily soil water dynamics, which provides a direct pathway for tracing how rainfall dataset differences propagate through phenology, biomass accumulation, and yield formation (Keating et al., 2003). APSIM has been previously validated for both study regions (Msongaleli et al., 2015; MacCarthy et al., 2021). Six widely used GPPs are integrated into APSIM and compared against simulations driven by local station data, with all non-climatic inputs held constant to isolate the influence of the rainfall dataset.

The objectives are threefold. First, to evaluate how GPPs represent rainfall characteristics most critical for rainfed agriculture, including seasonal totals, rainy days, onset timing, and drought occurrence, relative to local station data. Second, to examine how rainfall dataset differences propagate into APSIM simulations of maize phenology and yield formation by holding non-climatic inputs constant. Third, to assess how different GPPs alter the structure of nitrogen × planting-density yield response surfaces and the derived economically optimal input levels. These analyses are conducted under historical climate variability and are framed as sensitivity diagnostics of crop model outputs and derived economic indicators under alternative rainfall datasets. Together, these three objectives distinguish meteorological performance, crop-model propagation, and decision-support implications as related but conceptually distinct questions, addressed in sequence throughout the manuscript.

2. Materials and methods

2.1. Conceptual Framework

The conceptual framework (Fig. 1) outlines the sequence of data inputs, analytical steps, and benchmarking procedures used to evaluate the performance of GPPs for crop model applications. GPPs and station data are treated as alternative weather inputs and processed independently through APSIM, together with identical soil and crop management information, to isolate the influence of rainfall dataset choice on model behaviour. Rainfall diagnostics assess how well each dataset captures key rainfall indicators, while crop modelling and sensitivity analysis quantify how dataset differences shape simulated phenology, yield formation, and yield responses across nitrogen and planting-density options. Outputs from GPP-driven simulations are then benchmarked against station-based results to assess the consistency of simulated crop outputs and identify the conditions under which GPPs can serve as credible alternatives to station data.

2.2. Descriptions of the study area

This study examines agricultural conditions in Kaffrine, Senegal, and Kongwa, Tanzania - two semi-arid regions with distinct rainfall regimes, temperature patterns, soils, and farming systems (Table 1). Both areas exhibit considerable local variability in biophysical conditions, which shapes crop performance and management practices.

The contrasting rainfall regimes of Kaffrine and Kongwa shape sowing opportunities, nitrogen requirements, and attainable yield levels. Likewise, differences in soil fertility, texture, and water-holding capacity influence crop growth and management responses. These biophysical conditions provide a useful context for assessing how gridded precipitation datasets affect APSIM-simulated phenology, yield formation, and yield responses across nitrogen and planting-density options in diverse smallholder systems.

2.3. Weather data

Long-term (1983–2020) daily weather data were obtained from the Tanzania Meteorological Authority (TMA) for Kongwa and from the Agence Nationale de l'Aviation Civile et de la Météorologie (ANACIM) for Kaffrine. These high-quality datasets included rainfall and temperature variables, which are essential for analyzing rainfall patterns. The station data were pre-validated and quality-checked according to World Meteorological Organization (WMO) standards, ensuring completeness and reliability. For crop modelling, daily solar radiation was estimated from temperature data using the Donatelli–Campbell model (Mavromatis, 2008), a widely used approach in data-scarce regions due to limited direct radiation measurements.

In addition to station records, this study utilized six widely used gridded rainfall datasets derived from satellite observations, ground-based station data, and reanalysis products—CHIRPS, CPC, ERA5 Ag, MERRA-2, MSWEP, and TAMSAT. These datasets were selected for their daily temporal resolution, extensive spatial coverage across Africa, and compatibility with APSIM, allowing consistent evaluation of how different rainfall inputs influence simulated crop responses. Their inclusion supports a systematic comparison of rainfall representation across diverse agroecological conditions. The station and all six GPP daily rainfall records were complete across the 1983–2020 period, with no missing values requiring gap-filling or exclusion.

Although gridded datasets represent area-based averages and station records represent point measurements, direct comparison remains common in climatological and agricultural modelling studies (Angulo et al., 2013; Ensor and Robeson, 2008; Mankin et al., 2025; Mourtzinis et al., 2016). Such comparisons provide a practical basis for evaluating the suitability of gridded rainfall datasets for crop-model applications when interpreted in terms of spatial representativeness rather than exact

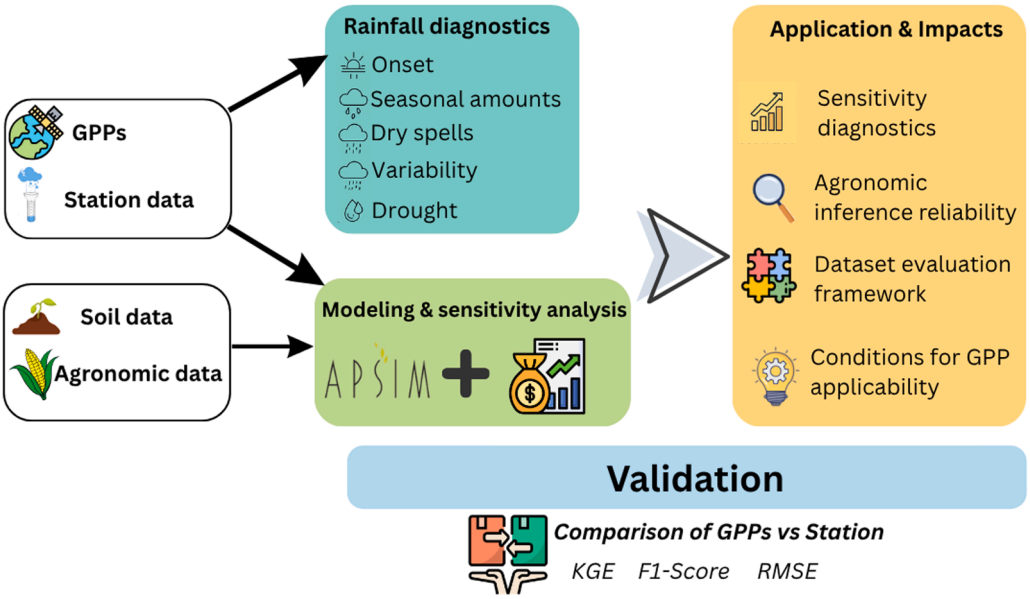


Fig. 1. Conceptual framework showing the use of GPPs and station data in APSIM simulations, combined with rainfall diagnostics and sensitivity analysis, followed by benchmarking against station-based results to assess the consistency of simulated crop outputs.

Table 1
Summary of the key agro-ecological characteristics of the study areas. Data summarized from FAO/IIASA Global Agro-Ecological Zones (GAEZ v4) (Fischer et al., 2021). Climatic values represent long-term site characteristics. Soil organic carbon and soil pH refer to the topsoil layer (0–30 cm).

Feature	Kaffrine, Senegal	Kongwa, Tanzania
Center coordinates	14°06′16″N, 15°32′59″W	6°12′00″S, 36°25′00″E
Elevation (m a.s.l.)	20 – 50 m	~1000 m
Annual rainfall (mm)	600 –700 mm (peak: Jul–Sep)	500 –700 mm (peak: Dec–Jan)
Growing season rainfall	450 – 650 mm (Jun–Sep)	400 – 600 mm (Dec–Apr)
Mean daily temperature	22–32°C (growing period), 28–38°C (annual)	18–27°C (growing period), 17–29°C (annual)
Rainfall pattern	Unimodal (Jun to early Oct)	Unimodal (Dec to Apr)
P/PET	0.12 – 0.59	0.17 – 0.51
Soil texture	Sandy loam	Sandy clay loam to clay
Drainage	Well-drained, low water holding	Freely drained
Soil organic carbon	0.5 – 1.3%	0.3 – 1.0%
Soil acidity	pH ~6.1 (slightly acidic)	pH ~5.9 (moderate acidic)
Farming system	Mixed crop-livestock (maize, millet, groundnut, ruminants)	Mixed crop-livestock (maize, sorghum, legumes, cattle, poultry)
Soil fertility management practices	Manure application, multiple fertilizer splits	Conservation agriculture, manure use, agroforestry

agreement. Because the GPPs differ in their native spatial resolution, no spatial resampling was performed. Instead, daily precipitation values were extracted at the station locations using a nearest-neighbour approach, thereby preserving each dataset at its native resolution. Nevertheless, results should be interpreted with awareness of the inherent scale mismatch between point observations and area-averaged gridded estimates. Table 2 summarizes the key characteristics of the GPPs used in this study.

2.4. Rainfall characteristics and variability

This study analyzed seasonal rainfall characteristics, interannual variability, and dry/wet year classification to support interpretation of APSIM-simulated crop responses. All analyses were applied consistently to station observations and the six GPP.

2.4.1. Seasonal rainfall patterns

Seasonal rainfall was characterized by using total rainfall, number of rainy days, rainfall intensity, and intra-seasonal variability. A rainy day was defined as any day receiving at least 1 mm of rainfall, consistent with the World Meteorological Organization (World Meteorological Organization, 2017) threshold. Temporal variability was quantified using the coefficient of variation (CV). Trends in seasonal rainfall and rainy-day counts were evaluated using the Mann–Kendall non-parametric test, which is widely used for detecting monotonic trends in hydro-climatic time series (Gocic and Trajkovic, 2013; Alemu and Dioha, 2020). These rainfall indicators were compared against station observations to assess how well the GPPs captured seasonal rainfall patterns.

2.4.2. Onset analysis

The timing of the rainy season was assessed using the Liebmann onset method, which identifies the start of sustained rainfall based on cumulative rainfall anomalies rather than a fixed rainfall threshold. This approach detects the transition to continuous rainfall and is widely applied in studies of seasonal rainfall dynamics in sub-Saharan Africa (Dunning et al., 2016; Joseph et al., 2023). Onset dates were computed for each year for both observed and gridded datasets. Using these onset dates, delayed onset was quantified using the mean standard deviation method to determine the delayed onset from the series of onset days (day of the year) for both observed and gridded datasets. The climatological period, i.e., 1991 – 2020 (as per WMO guideline, WMO, 2020), was used to compute both mean ($\bar{x}$) and standard deviation (SD). The delayed onset dates were defined as the onset dates that were later than the climatological mean onset date plus the standard deviation $D_{onset} > \bar{x} + SD$.

2.4.3. Drought classification

Drought conditions were characterized using MODIS-derived Normalized Difference Vegetation Index (NDVI) and seasonal rainfall totals. NDVI values were averaged over the rainy season for each year, and below-average vegetation cover was defined as NDVI values falling below the mean minus one standard deviation ($NDVI_{below} < \bar{x} - SD$). This threshold was used to identify years with reduced vegetation cover.

Table 2
Description of gridded precipitation datasets namely CHIRPS, CPC, MERRA2, MSWEP, and TAMSAT used in the current study acquired from their respective sources.

Dataset	Description	Resolution	Period	Primary Sources
CHIRPS	Provides daily and monthly rainfall estimates using thermal infrared data, passive microwave data, and gauge data (Du et al., 2023).	0.05° × 0.05°	1981–present	Satellite, ground stations
CPC	Provides global daily precipitation data based on gauge stations, interpolated to a 0.5° × 0.5° grid (NCAR, 2022).	0.5° × 0.5°	1979–present	Gauge stations
ERA5 Ag[a]	Provides daily meteorological data for agriculture and agro-ecological studies, aggregated from ERA5 hourly data and corrected to a finer topography using regression equations trained on ECMWF's high-resolution atmospheric model.	0.1° × 0.1°	1979–present	ERA5 reanalysis data, ECMWF HRES model
MERRA-2	A climate reanalysis dataset integrating satellite observations, ground-based measurements, and in situ data (Reichle et al., 2017).	0.5° × 0.625°	1980–present	Satellite, ground stations, in situ data
MSWEP	High-resolution global rainfall dataset combining satellite observations, ground-based measurements, and reanalysis data (Beck et al., 2019).	0.1° × 0.1°	1979–present	Satellite, ground stations, reanalysis data
TAMSAT	Provides detailed rainfall estimates and soil moisture forecasts for Africa, calibrated with ground-based measurements (Tarnavsky et al., 2014).	0.0375° × 0.0375°	1983–present	Satellite, ground stations

[a] In this manuscript, ERA5 Ag will be referred to simply as ERA5

Seasonal NDVI values were correlated with total rainfall to determine rainfall levels associated with below-average vegetation cover. From this relationship, rainfall thresholds indicative of dry seasons were identified as 400 mm for Kaffrine and 450 mm for Kongwa (see Figure S1 in the supplementary material S2). Seasonal conditions were classified as dry when total rainfall fell below these thresholds and wet when rainfall met or exceeded them. The same thresholds were applied to the GPP to evaluate their ability to reproduce observed drought classifications.

2.5. Modelling with observed precipitation and GPP datasets

2.5.1. APSIM model setup and input parameterization

APSIM (v7.10) was used as a process-based crop model to simulate maize growth and yield responses to rainfall variability under defined soil and management conditions. Its ability to represent genotype × environment × management interactions, simulate daily soil water dynamics, and capture crop development and yield formation under semi-arid and sub-humid African conditions has been widely demonstrated, relative to alternative process-based models such as DSSAT and AquaCrop (Keating et al., 2003; Msongaleli et al., 2015; MacCarthy et al., 2021). In particular, the explicit representation of soil water balance and crop growth processes allows rainfall variability to be translated into changes in water availability, crop development, and yield formation within the model. APSIM is therefore well suited for evaluating how differences in rainfall datasets propagate into crop simulations and influence simulated crop responses.

Agronomic and soil data collected in the respective study areas were used to parameterize APSIM simulations and define baseline conditions. Focus group discussions with local farmers and extension agents informed key management inputs, including sowing windows, fertilizer application rates and frequencies, and planting density. In each village, a mixed group of farmers representing the predominant farming systems was selected with the assistance of local extension agents and village leaders. Discussions followed a structured guide covering sowing dates, fertilizer management, planting density, and other practices relevant to APSIM parameterization. Information was recorded through detailed notes taken by the research team and synthesized to define the baseline management settings used in the simulations prior to the sensitivity analysis.

Soil samples were collected from representative fields in the study areas and analyzed at the International Institute of Tropical Agriculture (IITA) laboratory in Tanzania to determine physical and chemical properties, including texture, organic matter content, pH, and macronutrient levels. Soil hydraulic properties, including field capacity (DUL), permanent wilting point (LL15), and saturation (SAT), were estimated using the USDA Soil Water Calculator (Saxton and Rawls, 2006) and used to parameterize soil profiles in APSIM. Complete details of soil profiles are provided in Supplementary Tables S1 and S2.

2.5.2. Long-term simulations with station and GPP inputs

Long-term simulations were conducted using a single medium-maturing maize cultivar (*Dekalb XL82*) over the 1983–2020 period. *Dekalb XL82* was selected because it is commonly grown in semi-arid smallholder systems and has well-established APSIM calibration parameters for both study regions (Msongaleli et al., 2015; MacCarthy et al., 2021). Planting density (5 plants m^{-2}) and nitrogen rate (40 kg N ha^{-1}) were chosen to represent typical low-input management conditions in East and West African drylands, where fertilizer use is often below 20–40 kg N ha^{-1} and plant populations remain modest due to moisture and nutrient constraints (Sheahan and Barrett, 2016; Tittonell and Giller, 2012). Using representative low-input settings ensured that rainfall variability remained the dominant driver of simulated crop performance in this analysis. For each site, simulations were run separately with station rainfall and with each of the six GPP datasets, while all non-climatic inputs were held constant. This setup ensured that differences in simulated phenology, biomass accumulation, and grain yield reflected only differences in rainfall inputs. Because the GPPs are precipitation-only products, temperature, radiation, and CO2 were identical across all runs; only rainfall was varied.

Sowing windows were location-specific: November–January in Kongwa and June–July in Kaffrine. Sowing was triggered when $\geq$ 20 mm of rainfall accumulated within three consecutive days. Site-specific soil profiles were derived from local field surveys (Tables S1–S2). To avoid carryover effects across the 38-year simulation period, soil water and nitrogen were reset to the same baseline

conditions one month before the start of each growing season, effectively re-initializing the model independently for each season. Cultivar genetic parameters were adopted from previously published APSIM calibration studies for the respective regions (Msongaleli et al., 2015; MacCarthy et al., 2021), which calibrated crop development and grain-yield performance against local experimental observations. Model performance was evaluated using observed maize yields from 2014 to 2020 obtained from the Tanzania Agricultural Research Institute (TARI) for Kongwa and the Agence Nationale de Conseil Agricole et Rural (ANCAR) for Kaffrine. Validation statistics (e.g., R^2 and NRMSE) are presented in Supplementary Figure S2. Simulated yields showed good agreement with observed yields at both sites (Kaffrine: $R^2 = 0.79$, NRMSE = 0.35 t ha^{-1}; Kongwa: $R^2 = 0.81$, NRMSE = 0.15 t ha^{-1}), supporting use of the calibrated model for the long-term sensitivity simulations described below.

2.5.3. Sensitivity analysis of agronomic and economic responses to rainfall dataset choice

Building on the long-term simulations, a broader sensitivity analysis was conducted to assess the stability of simulated crop responses and economic indicators under rainfall dataset substitution across varying management conditions. This analysis extends the signal-propagation framework by systematically examining how alternative rainfall datasets influence yield response patterns across cultivars, nitrogen application, planting densities, and seasonal conditions, and whether these response structures remain consistent when translated into economic indicators. By expanding the genotype × management × seasonal decision space within a controlled simulation environment, the sensitivity analysis quantifies how rainfall datasets influence yield responses across nitrogen and planting density gradients. These effects were then evaluated with respect to economically optimal nitrogen rates and planting densities. In doing so, it provides a structured evaluation of the robustness of response patterns to rainfall dataset choice.

To generate the sensitivity dataset, a full factorial simulation design was implemented in APSIM, combining two maize cultivars (*Dekalb XL82* and *SC401*), location-specific sowing windows (three in Kongwa: November, December, January; and two in Kaffrine: June, July), seven nitrogen rates (0, 40, 60, 80, 100, 120, and 160 kg N ha^{-1}), and four planting densities (4, 5, 6, and 7 plants m^{-2}). Seasonal conditions were represented using wet and dry years classified according to the rainfall-based procedure described in Section 2.4. These factorial dimensions reflect the range of management intensities and cultivar choices commonly observed in semi-arid smallholder systems, while providing sufficient variability to detect how rainfall dataset differences propagate into agronomic and economic indicators.

Each factorial combination was simulated independently using station rainfall and each of the six GPP datasets. All non-climatic inputs, including soil profiles, sowing rules, and crop management logic, were held constant across simulations to ensure that differences in simulated outcomes arose solely from the interaction between rainfall datasets and the factorial treatment structure. Model outputs included grain yield, above ground biomass, and seasonal water balances at various crop phenological stages. These outputs provided a comprehensive basis for evaluating whether GPP rainfall inputs preserve the direction and magnitude of agronomic responses across contrasting genotype × management × seasonal environments.

To summarize the factorial results and enable comparison across rainfall datasets, yield response surfaces were calculated for each combination of cultivar, sowing window, and seasonal condition. Three functional forms were used (quadratic, quadratic plateau, and linear plateau) to capture nonlinear responses to nitrogen and planting density. Models were fitted using APSIM-simulated yields, and performance was assessed using R^2, RMSE, and Kling–Gupta Efficiency (KGE). Definitions and rationale for these metrics are provided in Section 2.6. Consistent with previous agronomic modelling studies, the quadratic model provided the best overall fit and was therefore selected for subsequent analysis (Schwalbert et al., 2018; Zhang et al., 2020). The fitted response surface took the form:

$$Y = a + bN + cN^2 + dD + eD^2 + fND \quad (1)$$

where Y is simulated grain yield (kg ha^{-1}), N is nitrogen rate (kg ha^{-1}), D is planting density (plants m^{-2}), and *a–f* are coefficients capturing main effects and interactions. This formulation (Eq. 1) enabled a consistent comparison of response curvature, marginal returns, and interaction structure across rainfall datasets. Response surfaces were fitted separately for each cultivar, rather than through an explicit interaction term, because the cultivars differ in underlying genetic coefficients governing phenology and radiation use efficiency, not in a simple shift of the nitrogen-density surface. This also mirrors on-farm decision-making, where cultivar is typically chosen first, with nitrogen rate and planting density optimized conditional on that choice.

Economic implications were evaluated by translating yield response surfaces into net-revenues. The net revenue (NR) was calculated as shown in Eq. 2 below:

$$NR = (Y \times Grain\ price) - (N \times Fertilizer\ cost) - (Seed\ requirement \times Seed\ price) \quad (2)$$

Seed requirement was calculated based on planting density and average seed weight. Prices were held at a single representative value from 2025 for each location rather than varying by year, so that variation in simulated net revenue reflects differences in rainfall dataset and resulting yield rather than confounding market price variability. For Kongwa, grain and seed prices were obtained from the Tanzanian Ministry of Agriculture (2025). For Kaffrine, prices were sourced from the Selina Wamucii (2025). Data of fertilizer costs were obtained from the International Fertilizer Development Center's Africa Fertilizer dashboard (IFDC, 2025). Exchange rates were obtained from the Bank of Tanzania and XE.com for consistency across currencies.

2.6. Performance evaluation metrics

To evaluate the performance of GPPs, a suite of statistical metrics was used, distinct from those applied in Section 2.5 for agronomic response surface fitting. In this section, the focus is on assessing rainfall dataset quality relative to station observations. For continuous evaluation, metrics included the Kling-Gupta Efficiency (KGE), coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Among these, KGE was selected as a primary performance indicator due to its ability to integrate correlation, bias, and variability into a single measure (Mathevet et al., 2023). This composite nature makes it particularly effective for assessing how well GPPs replicate station-based rainfall data, as supported by previous studies (e.g., Alfieri et al., 2022; Beck et al., 2019).

For classification tasks, the F1 score was used, with 1.0 as perfect agreement. This metric was applied in two distinct contexts: classifying delayed versus non-delayed rainfall onset (Section 3.1.2) and classifying daily rainfall intensity categories against station-observed classes (Sections 3.1.1 and 3.1.3); in both cases, it summarizes agreement between GPP- and station-based binary classifications for the category of interest. This metric balances precision and recall, offering a reliable assessment of classifier performance under imbalanced class distributions (Riyanto et al., 2023). The mathematical definitions of all performance metrics are provided in Supplementary Section S2. All data processing and statistical analyses were performed in Python.

3. Results

3.1. Representation of seasonal rainfall characteristics

3.1.1. Rainfall distribution, variability, and trends

Probability of exceedance distributions for seasonal rainfall totals

and rainy-day frequency in Kaffrine and Kongwa were generally aligned with the relative ranking of Kling–Gupta Efficiency (KGE) values, although differences across sites and variables were evident (Fig. 2).

Seasonal rainfall totals in Kaffrine (Fig. 2a) showed that most products followed the station distribution across central quantiles, with divergence emerging toward the upper tail. CHIRPS, MSWEP, and MERRA2 recorded comparable KGE values (≈ 0.51–0.58). CHIRPS and MSWEP tracked the station exceedance curve closely across most quantiles, whereas MERRA2 exhibited a flatter distribution with a slightly extended upper tail relative to the station record. CPC and ERA5 displayed lower agreement (KGE ≤ 0.37) and showed displacement toward lower rainfall amounts at higher exceedance probabilities. Seasonal totals in Kongwa (Fig. 2c) exhibited generally stronger agreement than in Kaffrine. CHIRPS, MSWEP, and TAMSAT recorded KGE values ≥ 0.78 and closely followed the station distribution across central and upper quantiles. In contrast, ERA5 and MERRA2 showed substantially lower KGE values (≈ 0.26–0.29), with ERA5 displaced toward lower totals and MERRA2 exhibiting elevated values in the upper tail of the distribution.

Rainy-day frequency in Kaffrine (Fig. 2b) exhibited greater spread among products than seasonal totals. CHIRPS recorded KGE = 0.75, CPC, TAMSAT, and ERA5 recorded moderate values (≈ 0.47–0.50), MSWEP recorded 0.37, and MERRA2 recorded −0.65. The exceedance curve of MERRA2 was shifted upward relative to the station, indicating higher rainy-day counts across much of the distribution. Rainy-day frequency in Kongwa (Fig. 2d) showed limited agreement across most products. CPC (0.55) and CHIRPS (0.50) exceeded a KGE of 0.50, while MERRA2, MSWEP, and TAMSAT displayed upward displacement of exceedance curves relative to the station distribution.

Daily rainfall intensity was evaluated using the categories of Wang et al. (2021): dry days (<1 mm), light rain (<5 mm), moderate rain (5–20 mm), heavy rain (20–50 mm), and very heavy rain (>50 mm). Across both sites, F1 scores for dry days ranged from 0.51 to 0.82. F1 values were lower for higher intensity classes, with most products recording reduced detection rates for moderate, heavy, and very heavy rainfall events. TAMSAT, MSWEP, ERA5, and CHIRPS recorded comparatively higher F1 values for moderate rainfall relative to other products, whereas CPC and MERRA2 recorded lower values across intensity classes beyond dry days. Full F1 scores for rainfall intensity detection are provided in Supplementary Section S4.2 (Tables S4–S5).

Trend analysis using the Mann–Kendall test indicated partial correspondence with station records. In Kaffrine, CPC detected significant increasing trends in both seasonal rainfall totals and rainy-day frequency, matching the station results, while other products did not detect significant trends. In Kongwa, most detected trends were statistically insignificant. MERRA2 detected a significant increase in seasonal rainfall, and CPC detected a significant increase in rainy-day frequency. Complete Mann–Kendall trend results are provided in Supplementary Section S4.1 (Table S3).

3.1.2. Rainfall onset variability

Rainfall onset variability was evaluated using KGE, and delayed onset events were assessed using the F1 score (Fig. 3). In Kaffrine, CHIRPS achieved the highest onset variability skill (KGE = 0.53), while

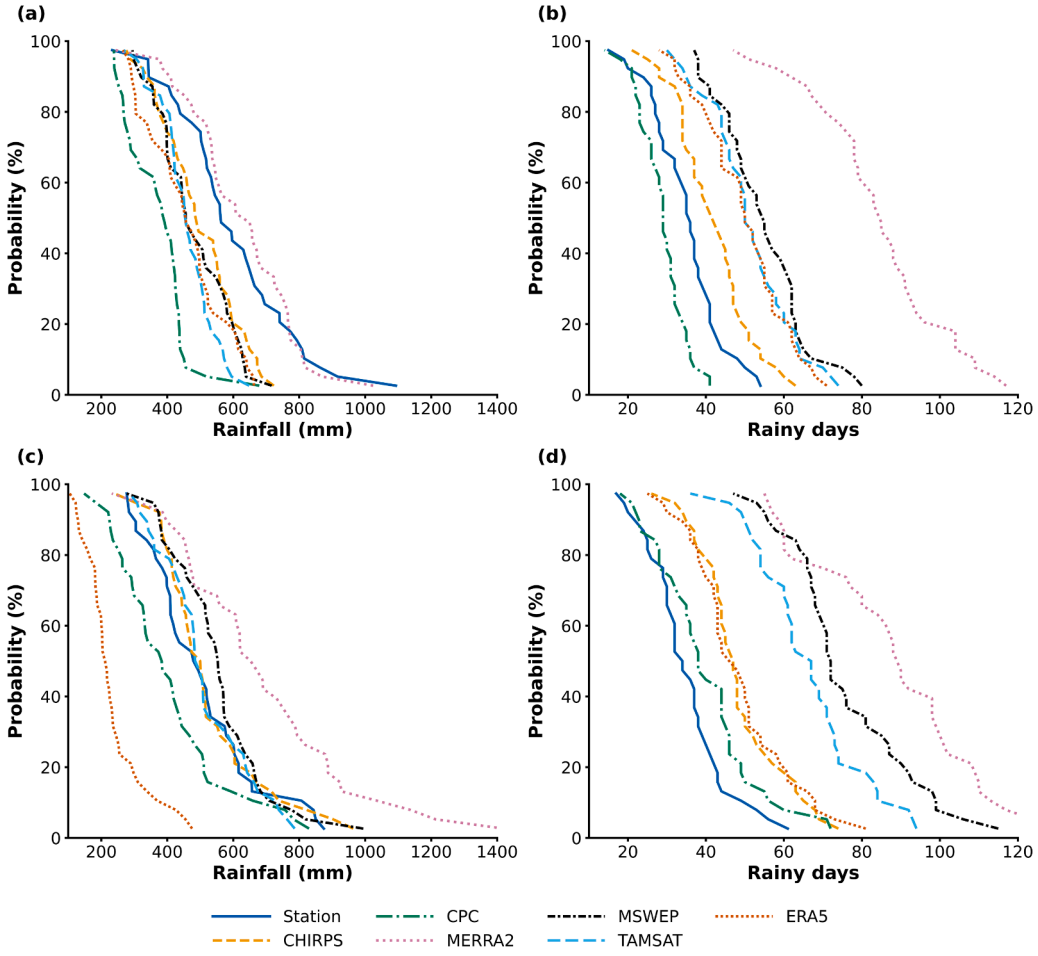


Fig. 2. Probability of exceedance charts on seasonal rainfall amounts and the number of rainy days in two regions: Kaffrine (panels a and b) and Kongwa (panels c and d). Seasonal rainfall amount refers to total accumulated rainfall over the growing season (Table 1); a rainy day is defined as any day with ≥ 1 mm rainfall (Section 2.4.1). GPP abbreviations: CHIRPS, CPC, ERA5, MERRA-2, MSWEP, TAMSAT (Table 2).

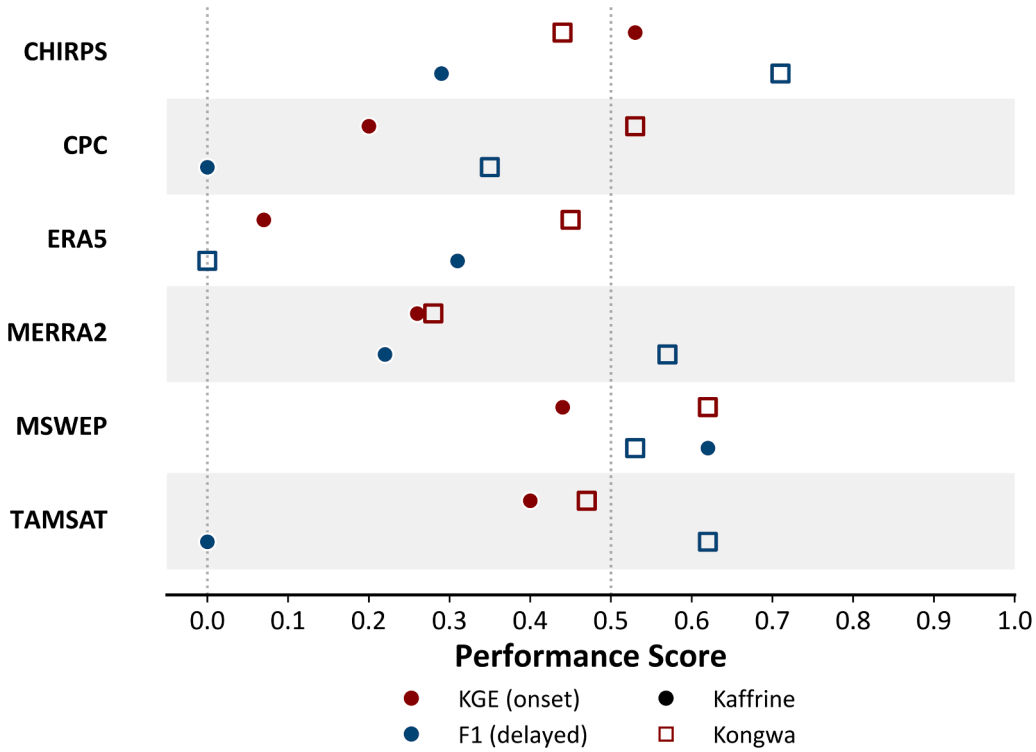


Fig. 3. Performance of gridded precipitation datasets in representing rainfall onset indicators at Kaffrine and Kongwa. KGE evaluates agreement in onset timing (continuous variable), while the F1 score assesses delayed onset detection (binary classification). Circles denote Kaffrine and squares denote Kongwa; blue markers represent F1 scores and red markers represent KGE. Vertical dotted lines mark scores of 0.0 (no skill) and 0.5 (moderate skill).

ERA5 showed the lowest performance (0.07). MSWEP (0.44) and TAMSAT (0.40) clustered around 0.40–0.45, whereas CPC (0.20) and MERRA2 (0.26) showed lower skill. For delayed onset detection, MSWEP achieved the highest performance (F1 = 0.62), while CPC and TAMSAT did not detect delayed events (F1 = 0.00). In Kongwa, MSWEP (0.62) and CPC (0.53) exceeded 0.50, while CHIRPS, ERA5, and TAMSAT clustered near 0.45. For delayed onset detection, CHIRPS (0.71) and TAMSAT (0.62) achieved the highest performance, followed by MERRA2 (0.57) and MSWEP (0.53), whereas ERA5 recorded no successful detection (F1 = 0.00). Complete values are provided in Fig. 3 and Supplementary Table S3.

3.1.3. Drought occurrence detection

Drought detection performance was evaluated using the F1 score, with drought seasons defined as years in which seasonal rainfall fell below 400 mm in Kaffrine and 450 mm in Kongwa. Fig. 3 shows the distribution of F1 scores across products for both sites. In Kaffrine, F1 scores ranged from 0.40 to 0.60. CHIRPS and ERA5 achieved the highest values (0.60), followed closely by MERRA2 (0.59) and MSWEP (0.58). TAMSAT recorded 0.51, while CPC showed the lowest score (0.40). Differences among products were relatively narrow, with four datasets clustering between 0.58 and 0.60. In Kongwa, F1 scores ranged from 0.36 to 0.79. TAMSAT achieved the highest value (0.79), followed by CPC (0.74) and CHIRPS (0.73). MSWEP showed intermediate skill (0.56), while ERA5 (0.40) and MERRA2 (0.36) recorded lower values. The spread among products was larger in Kongwa than in Kaffrine, reflecting greater inter-product variability in drought classification.

Across rainfall characteristics, product performance varied systematically by metric and site. Agreement with station data was generally higher for seasonal rainfall totals than for rainfall onset detection, particularly in Kaffrine where onset skill remained below 0.50 for most datasets. Detection of rainfall intensities showed relatively strong performance for dry days but substantially lower F1 scores for moderate and heavy rainfall classes. Drought classification skill exhibited narrower inter-product differences in Kaffrine but wider dispersion in Kongwa. No single rainfall product consistently ranked highest across rainfall totals, onset timing, intensity detection, and drought occurrence. These metric- and site-specific rainfall differences provide the basis for evaluating how rainfall dataset substitution propagates into simulated crop performance and model-derived outputs (Section 3.2).

3.2. Propagation of rainfall dataset differences into APSIM-simulated crop responses

This section presents how differences among rainfall datasets propagated through APSIM simulations and influenced simulated crop responses, including drought impacts, sowing-window performance, cultivar response patterns, estimated input optima, and yield variability relative to station-based simulations.

3.2.1. Simulated yield responses and within-season alignment across rainfall datasets

In Kaffrine (Fig. 4a), yields simulated using the station dataset exhibited a relatively narrow interquartile range compared with Kongwa. Across GPPs, median yields were generally comparable to the station baseline, although dispersion varied among datasets. CPC displayed a wider spread toward lower yields, while TAMSAT showed a slightly higher median relative to the station. In Kongwa (Fig. 4b), yield variability was substantially greater across all datasets. ERA5 produced markedly lower median yields relative to the station, whereas MERRA2 generated higher medians and a broader upper distribution. Other products (CHIRPS, CPC, MSWEP, and TAMSAT) showed distributions that overlapped with the station baseline but differed in spread.

Kolmogorov–Smirnov comparisons between GPP-based and station-derived distributions indicated that in Kaffrine, MERRA2 ($p = 0.30$) and ERA5 ($p = 0.41$) did not differ significantly from the station, whereas CHIRPS, CPC, MSWEP, and TAMSAT showed significant differences ($p < 0.05$). In Kongwa, the pattern reversed: MERRA2 and ERA5

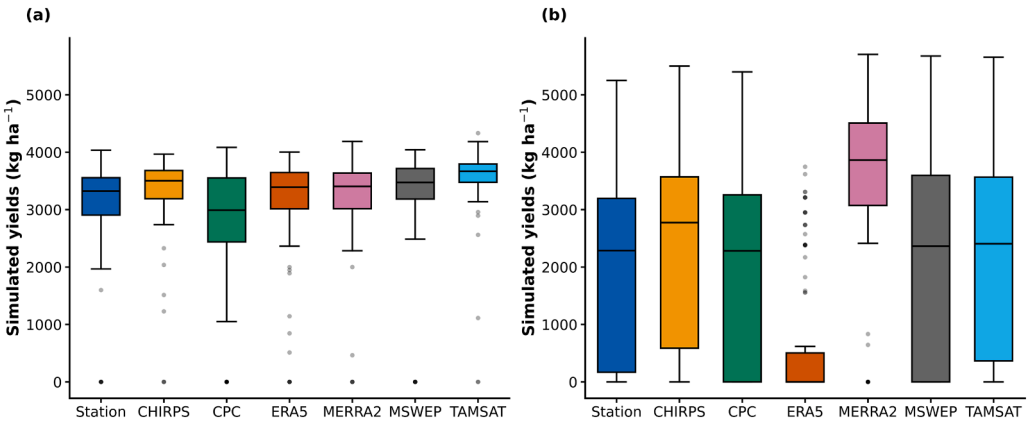


Fig. 4. Comparisons of simulated yield distributions using GPPs and the station dataset, and among different GPPs, for (a) Kaffrine and (b) Kongwa. All distributions, including the "Station" category, represent APSIM-simulated yields. The station category reflects simulations driven by station-recorded rainfall rather than independently observed field yields.

differed significantly from the station distribution, while CHIRPS, CPC, MSWEP, and TAMSAT did not ($p > 0.05$). Yield agreement (KGE) showed that in Kongwa, CHIRPS (0.60), TAMSAT (0.58), and MSWEP (0.53) had the highest agreement with station-based yields, whereas MERRA2 (0.01) and ERA5 (−0.24) showed low agreement. In Kaffrine, KGE values were lower overall, with CPC (0.35), TAMSAT (0.26), and MSWEP (0.24) ranking highest among the evaluated datasets.

3.2.2. Within-season alignment of rainfall phases and crop-state variables

Across both sites, correlations between simulated yield and pre-anthesis rainfall were weak and inconsistent. In Kaffrine, correlations ranged from −0.32 (MERRA2) to −0.01 (CHIRPS and TAMSAT) (Figure S3). In Kongwa, values ranged from −0.10 (CPC) to 0.19 (MERRA2) (Figure S4), with all coefficients remaining close to zero. In contrast, correlations between simulated yield and post-anthesis rainfall were consistently higher across most datasets. In Kaffrine, yield–post-anthesis correlations ranged from 0.20 (station) to 0.69 (CPC), with moderate values for MSWEP (0.51), MERRA2 (0.46), and TAMSAT (0.38). In Kongwa, correlations ranged from 0.33 (CHIRPS) to 0.69 (CPC), with moderate values for MERRA2 (0.55) and MSWEP (0.51).

Comparisons of post-anthesis cumulative rainfall against the station baseline showed limited significant deviations in Kaffrine (Table S6). Only CPC differed significantly (mean difference = 107 mm, $p < 0.001$), while CHIRPS, MERRA2, MSWEP, and TAMSAT did not differ significantly. Post-anthesis biomass accumulation showed no significant differences, while grain filling rate differed significantly only for TAMSAT ($p = 0.020$), and soil moisture differed for CPC and MERRA2. In Kongwa, significant post-anthesis deviations were observed only for MERRA2 (Table S7). Cumulative post-anthesis rainfall differed significantly from the station baseline (mean difference = −170 mm, $p < 0.001$). MERRA2 also showed significant differences in post-anthesis biomass accumulation, grain filling rate, and soil moisture. Other datasets did not differ significantly from the station baseline for these variables.

Across both sites, correlations between yield and post-anthesis rainfall exceeded those for pre-anthesis rainfall. Statistically significant deviations in post-anthesis rainfall and associated crop-state variables were limited to specific datasets rather than observed uniformly across products.

3.2.3. Sensitivity of simulated yield responses to rainfall dataset choice

Simulated yield responses differed across rainfall datasets and management combinations, with differences in the magnitude of deviations and in the location of response-surface optima relative to the station baseline (Figures S6–S8; Tables S8a–S8b). In Kaffrine, deviations in economically optimal nitrogen rate (EONR) relative to the station baseline were generally small across rainfall datasets (Figure S7; Table S8a). For CHIRPS, MSWEP, TAMSAT, and MERRA2, EONR differences were typically within ±10 kg N ha^{-1} across climate–cultivar combinations. CPC showed moderate downward shifts under selected combinations, while ERA5 exhibited larger deviations but remained within the same order of magnitude. Plant density (EOD) deviations were limited, usually within ±1 plant m^{-2} across datasets. Net revenue deviations were modest in Kaffrine, with TAMSAT generating the largest positive differences under dry conditions, while CPC and ERA5 produced the largest negative differences under wet conditions. No rainfall dataset showed systematic divergence across all factorial combinations.

In Kongwa, deviations were larger and more dataset-specific (Figure S8; Table S8b). Under wet conditions, CHIRPS, MSWEP, and TAMSAT produced reductions in EONR relative to the station baseline, in some cases exceeding −30 to −40 kg N ha^{-1} . CPC showed smaller reductions. Under ERA5, the simulated crop failed to reach maturity in wet-condition scenarios (Table S8b). Under dry conditions, most rainfall datasets increased EONR relative to the station baseline, with CPC, MSWEP, and TAMSAT frequently exceeding + 25 kg N ha^{-1} for selected combinations. EOD deviations in Kongwa remained comparatively small (generally 0 to +3 plants m^{-2}), although variability across climate states was evident. Net revenue deviations displayed larger amplitude than in Kaffrine. Wet-season simulations produced substantial negative deviations for MSWEP and CHIRPS, whereas dry-season simulations generated large positive deviations for CPC and TAMSAT under selected combinations.

4. Discussion

This study demonstrates that rainfall dataset evaluation for crop modelling should extend beyond agreement in rainfall indicators to include the stability of model-derived outputs under dataset substitution. By explicitly linking rainfall diagnostics with nitrogen × planting-density yield response surfaces and derived economic indicators, the analysis provides a structured basis for assessing whether rainfall datasets can be used interchangeably in crop modelling applications, offering a more comprehensive basis for assessing GPP suitability than approaches based on rainfall statistics or single-output crop-model comparisons alone.

Such diagnostics complement, rather than replace, conventional rainfall validation. In this sense, a rainfall dataset's meteorological agreement with station observations is conceptually distinct from its decision-support reliability, since a dataset that closely tracks station rainfall statistically does not necessarily lead to the same agronomic or economic recommendation as station data.

4.1. Rainfall representation across metrics and sites

Rainfall dataset performance was highly dependent on both the indicator considered and the local context. No dataset consistently outperformed others across all rainfall characteristics, echoing earlier evaluations that emphasize GPP's regional variability rather than their universal transferability (Satgé et al., 2020; Cao et al., 2023). CHIRPS, TAMSAT, MSWEP, and CPC often aligned more closely with station observations, whereas ERA5 and MERRA2 showed larger deviations in Kongwa than in Kaffrine, underscoring the influence of site-specific rainfall regimes on dataset reliability.

Beyond differences in seasonal totals, rainfall datasets showed systematic discrepancies in event frequency and intensity. Most gridded products reported a higher number of rainy days than station observations. This tendency reflects the spatial averaging inherent in gridded datasets, which smooths localized rainfall extremes and redistributes precipitation into lighter but more frequent events across grid cells. Such smoothing inflates low-intensity rainfall while attenuating heavy rainfall signals, highlighting the structural contrast between area-based grid estimates and point-based station measurements. Similar limitations in intensity representation have been reported for satellite-derived rainfall products (Cao et al., 2023). In semi-arid systems, where short dry spells and episodic rainfall events strongly regulate crop establishment and early growth, these structural differences can influence simulated yield variability more than moderate discrepancies in seasonal rainfall totals.

Trend detection highlighted another dimension of dataset performance. While some products captured interannual fluctuations, they did not consistently reproduce longer-term tendencies. This is in line with previous evaluations that show that agreement in variability does not necessarily translate into reliable trend representation, due to inhomogeneity in station inputs and algorithmic biases (Dinku, 2019; Satgé et al., 2020). In our study, discrepancies were more pronounced in Kongwa, reinforcing the importance of local rainfall regimes. Reliable trend representation is therefore critical for agricultural applications, as it shapes interpretation of long-term climate signals alongside seasonal variability.

Regional contrasts were particularly evident for reanalysis products. ERA5 and MERRA2 showed comparatively stronger agreement in Kaffrine than in Kongwa. Previous studies have documented that reanalysis performance is influenced by assimilated station density and regional data availability (Reichle et al., 2017; Dinku, 2019; Ahmed et al., 2024). Although station density was not quantified in this study, the stronger agreement observed in Kaffrine may reflect differences in station coverage and data assimilation quality between West Africa and East Africa, which can affect how well reanalysis products capture local rainfall variability.

4.2. Propagation of rainfall differences into simulated yield

Rainfall dataset differences resulted in contrasting patterns of simulated yield variability between the two case studies. In Kaffrine, yield distributions across datasets remained relatively close to the station reference, with limited variability among products. In contrast, in Kongwa, simulated yields showed substantially greater dispersion and divergence across datasets. This contrast reflects Kongwa's greater exposure to intermittent dry spells during crop establishment relative to Kaffrine's more consistent early-season rainfall, a pattern consistent with previous studies showing that crop model sensitivity to weather inputs varies across agro-climatic settings (Angulo et al., 2013; Hoffmann et al., 2017).

These contrasts reflect differences in cropping calendars and seasonal rainfall structure. In Kaffrine, sowing is concentrated in June–July and is typically followed by increasing rainfall, resulting in relatively stable conditions during establishment and early growth. Under such conditions, similarities in early-season moisture across rainfall datasets lead to consistent yield formation despite differences in rainfall statistics. In contrast, sowing in Kongwa spans November–January and frequently coincides with intermittent dry spells. As a result, small differences in rainfall representation during establishment alter early crop development and propagate into larger yield differences, a mechanism also reported for semi-arid cropping systems (MacCarthy et al., 2021; Akinseye et al., 2016). The greater yield dispersion in Kongwa therefore reflects sensitivity to early-season moisture variability rather than seasonal rainfall totals.

Within-season diagnostics indicated that simulated yield was more strongly influenced by post-anthesis rainfall than by pre-anthesis rainfall at both sites, consistent with the established importance of reproductive-stage water availability for grain production (Rötter and Van Keulen, 1997; Frieler et al., 2017; MacCarthy et al., 2021). However, stronger alignment with post-anthesis rainfall did not translate into closer agreement among rainfall-driven crop simulations.

The largest yield differences across datasets were associated with variability during crop establishment, particularly in Kongwa where sowing windows expose crops to intermittent dry spells. Early-season rainfall differences influenced initial crop development and led to divergent crop states prior to the reproductive stage, such that even under similar post-anthesis conditions, yield outcomes remained different. These differences were subsequently propagated through crop growth processes represented in APSIM, whereby differences in early-season water availability influenced crop growth trajectories, biomass accumulation, and ultimately grain yield formation (Keating et al., 2003; Boote et al., 2021). In this study, however, post-anthesis rainfall largely governed overall yield levels, whereas early-season rainfall distribution determined divergence in crop growth trajectories among rainfall-driven simulations.

Overall, it was not the total seasonal rainfall alone, but rainfall distribution within the growing season that determined the stability of crop model outcomes. In environments with more reliable rainy seasons, alternative rainfall datasets produced similar simulated yields, whereas in environments with variable early-season moisture the same rainfall differences resulted in substantial divergence. This aligns with previous studies showing that crop model sensitivity to weather inputs is governed by phenological timing and stress occurrence rather than aggregate rainfall totals (Challinor et al., 2014; Stern and Cooper, 2011). Assessing rainfall datasets for crop modelling applications therefore requires evaluating the sensitivity of simulated outcomes in addition to meteorological agreement.

4.3. Sensitivity of agronomic response surfaces to rainfall dataset choice

Testing this sensitivity across a full factorial design spanning cultivar, sowing window, nitrogen rate, and planting density, rather than a single management scenario, provides a methodologically robust basis for evaluating whether rainfall dataset differences matter consistently or only under specific management conditions.

Rainfall dataset substitution affected the magnitude of simulated optimal input levels differently across environments. In Kaffrine, economically optimal nitrogen rates (EONR) and planting densities (EOD) remained relatively stable across datasets, whereas in Kongwa, optimal nitrogen levels and associated net revenue varied more substantially. In some cases, these differences were large enough to alter the recommended nitrogen rate (Section 3.2.3). This contrast indicates that the sensitivity of model-derived outputs to rainfall dataset choice is context-dependent, with stronger effects in environments characterized by higher rainfall variability; therefore, model responses should not be assumed to be uniformly stable across rainfall datasets (Challinor et al., 2014).

Rainfall dataset substitution altered the magnitude of optimal input levels while maintaining broadly similar yield response patterns. Across datasets, nitrogen–density relationships remained unimodal and cultivar-specific responses were preserved, even where optimal input levels shifted. These results indicate that rainfall dataset differences

primarily affect the location and magnitude of simulated optima, rather than the underlying response relationships represented in the model.

The stronger sensitivity observed in Kongwa likely reflects the interaction between nitrogen response and variable water availability, particularly where soil moisture conditions fluctuate more strongly across seasons and sowing windows (Akinseye et al., 2016). Under such conditions, rainfall dataset choice influences not only simulated yield levels but also the economic interpretation of input use. This indicates that rainfall datasets producing similar yield patterns may still differ in the optimal input levels and net revenues they imply, consistent with findings that economically optimal input levels derived from crop models can vary substantially with environmental conditions and input–yield relationships (Schwalbert et al., 2018). For crop modelling applications, dataset evaluation should therefore consider not only agreement in simulated yields but also the stability of derived economic indicators across rainfall datasets.

4.4. Contribution and implications for climate–crop modelling

This study's contribution lies in demonstrating that different components of crop response are governed by distinct aspects of rainfall variability, rather than by dataset performance in aggregate. Simulated yield levels were comparatively stable across datasets, reflecting the dominant role of reproductive-stage water availability, whereas outputs sensitive to seasonal structure, such as sowing-window performance and optimal input levels, were more strongly affected by differences in early-season rainfall representation, particularly under intermittent dry spells and variable soil moisture during crop establishment. For climate-crop modelling applications, this implies that rainfall dataset suitability should be assessed at the level of the specific output of interest, rather than assumed uniform across an application.

This suggests that GPP suitability should be judged against the rainfall characteristics most relevant to the intended application. Where crop establishment occurs under a relatively reliable onset and consistent early-season moisture, as in Kaffrine, rainfall dataset substitution resulted in comparatively limited differences in simulated outcomes. In contrast, where establishment coincides with intermittent dry spells, as in Kongwa, differences in a dataset's representation of onset timing and early-season rainfall can translate into larger differences in downstream agronomic and economic outcomes. Consistent with the region-specific nature of GPP performance noted in Section 4.5, it is this evaluation framework, rather than the specific dataset rankings reported here, that is expected to generalize to other locations.

4.5. Limitations and scope

This study is based on two semi-arid environments, and the findings should therefore be interpreted within this domain. The results should not be used to select a specific rainfall dataset for other regions, as station density, product construction, and representation of meteorological processes vary across regions. This analysis is also specific to maize and should not be assumed to generalize to other crops.

The analysis is based on a single crop model configuration; structural uncertainty associated with alternative model formulations or parameterizations was not explicitly assessed and may influence the magnitude of simulated responses. The quantitative sensitivity reported here is therefore specific to APSIM and should not be assumed to hold under other crop modelling frameworks. Because temperature, radiation, and CO_2 were held constant to isolate the effect of rainfall dataset choice, this design does not capture how rainfall-driven differences might interact with alternative temperature or radiation datasets.

Rainfall datasets were evaluated at the scale of available station observations, and sub-grid heterogeneity was not explicitly resolved (Josephson et al., 2026). Net revenue estimates are based on fixed, single-year prices rather than a historical price series and do not capture market or price risk; reported economic indicators should be interpreted as rainfall-driven sensitivity rather than full historical profitability. Within these constraints, the results provide a consistent basis for assessing how rainfall dataset differences influence simulated crop responses under semi-arid conditions.

5. Conclusion

This study examined how substituting rainfall datasets influences simulated crop responses in process-based crop modelling across two semi-arid maize systems. The results show that rainfall dataset choice is context-dependent and that agreement in rainfall indicators alone does not ensure consistency in crop model outputs. Across the two study regions, dataset substitution led to limited differences in Kaffrine but substantially larger divergence in Kongwa, particularly in terms of yield variability and derived optimal input levels. These findings highlight that the suitability of rainfall datasets for agricultural applications depends on their influence on model-derived outcomes rather than on statistical rainfall performance alone. This has important implications for both research and project implementation, suggesting that climate data selection should be guided by its relevance for impact modelling.

By systematically testing rainfall dataset sensitivity across a full factorial design of cultivars, sowing windows, nitrogen rates, and planting densities, rather than a single management scenario, this study provides a robust evidence base for climate-risk assessment and agricultural decision-support in data-scarce regions. Although the analysis is based on two semi-arid maize systems and a single modelling framework, the results provide a basis for assessing rainfall dataset substitution in similar rainfed environments. Further work is needed to evaluate the extent to which these findings apply across additional crops, climates, and modelling approaches.

CRediT authorship contribution statement

Jacob Emanuel Joseph: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Data curation, Conceptualization. **Christoph Gornott:** Writing – review & editing, Visualization, Validation, Supervision, Data curation, Conceptualization. **Rahel Laudien:** Writing – review & editing, Validation, Methodology, Conceptualization. **Stefan Siebert:** Writing – review & editing, Validation, Supervision. **Omonlola Nadine Worou:** Writing – review & editing, Supervision, Data curation. **Anthony M. Whitbread:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Reimund P. Rötter:** Writing – review & editing, Validation, Supervision, Resources, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the Academy for International Agricultural Research (ACINAR), commissioned by the German Federal Ministry for Economic Cooperation and Development (BMZ), and implemented by ATSAF e.V. on behalf of the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH. We also gratefully acknowledge the support of the Tropical Plant Production and Agricultural Systems Modelling (TROPAGS) division in the Department of Crop Sciences at the University of Göttingen, Germany. Additionally, we extend our appreciation to the International Livestock Research Institute (ILRI) for funding the contributions of Jacob Emanuel Joseph and Anthony M. Whitbread to this study.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eja.2026.128344.

Data availability

Data will be made available on request.

References

Ahmed, J.S., Buizza, R., Dell'Acqua, M., Demissie, T., Pè, M.E., 2024. Evaluation of ERA5 and CHIRPS rainfall estimates against observations across Ethiopia. Meteorol. Atmos. Phys. 136 (3). https://doi.org/10.1007/s00703-024-01008-0.

Akinseye, F., Adam, M., Agele, S., Hoffmann, M., Traore, P., Whitbread, A., 2016. Assessing crop model improvements through comparison of sorghum (Sorghum bicolor L. moench) simulation models: A case study of West African varieties. Field Crops Res. 201, 19–31. https://doi.org/10.1016/j.fcr.2016.10.015.

Alemu, Z.A., Dioha, M.O., 2020. Climate change and trend analysis of temperature: the case of Addis Ababa, Ethiopia. Environ. Syst. Res. 9 (1), 27. https://doi.org/10.1186/s40068-020-00190-5. https://environmentalsystemsresearch.springeropen.com/articles/10.1186/s40068-020-00190-5.

Alfieri, L., Avanzi, F., Delogu, F., Gabellani, S., Bruno, G., Campo, L., Libertino, A., Massari, C., Tarpanelli, A., Rains, D., Miralles, D.G., Quast, R., Vreugdenhil, M., Wu, H., Brocca, L., 2022. High-resolution satellite products improve hydrological modeling in northern Italy. Hydrol. Earth Syst. Sci. 26 (14), 3921–3939. https://doi.org/10.5194/hess-26-3921-2022.

Alimagham, S., Van Loon, M.P., Ramirez-Villegas, J., Adjei-Nsiah, S., Baijukya, F., Bala, A., Chikowo, R., Silva, J.V., Soulé, A.M., Taulya, G., Tenorio, F.A., Tesfaye, K., Van Ittersum, M.K., 2024. Climate change impact and adaptation of rainfed cereal crops in sub-Saharan Africa. Eur. J. Agron. 155, 127137. https://doi.org/10.1016/j.eja.2024.127137.

Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P., Ewert, F., 2013. Characteristic 'fingerprints' of crop model responses to weather input data at different spatial resolutions. Eur. J. Agron. 49, 104–114. https://doi.org/10.1016/j.eja.2013.04.003.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2019. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. Bull. Am. Meteorol. Soc. 100 (3), 473–500. https://doi.org/10.1175/bams-d-17-0138.1.

Boote, K.J., Adam, M., Ahmad, I., Ahmad, S., Cammarano, D., Chattha, A.A., Claessens, L., Dimes, J., Durand, W., Freduah, B.S., Gummadi, S., Hargreaves, J., Hoogenboom, G., Tui, S.H., Jones, J.W., Khaliq, T., MacCarthy, D.S., Masikati, P., McDermid, S., Wajid, S.A., 2021. Understanding differences in climate sensitivity simulations of APSIM and DSSAT crop models. World Scientific (Europe) eBooks, pp. 15–46. https://doi.org/10.1142/9781786348791_0002.

Cao, W., Nie, S., Ma, L., Zhao, L., 2023. Statistical evaluation of the performance of gridded daily precipitation products from reanalysis data, satellite estimates, and merged analyses over global land. Remote. Sens. 15 (18), 4602. https://doi.org/10.3390/rs15184602.

Challinor, A.J., Watson, J., Lobell, D.B., Howden, S.M., Smith, D.R., Chhetri, N., 2014. A meta-analysis of crop yield under climate change and adaptation. Nat. Clim. Change 4 (4), 287–291. https://doi.org/10.1038/nclimate2153. https://www.nature.com/articles/nclimate2153.

Chatterjee, D., Singh, D., Singh, P.K., Fohrer, N., Singh, B.B., 2023. Performance evaluation of different gridded precipitation and CMIP6 model products with gauge observations for assessing rainfall variability under the historical and future climate change scenario over a semi-arid catchment, India. Phys. Chem. Earth/Phys. Chem. Earth, Parts A/B/C 131, 103433. https://doi.org/10.1016/j.pce.2023.103433.

Deutsch, C.A., Tewksbury, J.J., Tigchelaar, M., Battisti, D.S., Merrill, S.C., Huey, R.B., Naylor, R.L., 2018. Increase in crop losses to insect pests in a warming climate. Science 361 (6405), 916–919. https://doi.org/10.1126/science.aat3466.

Dinku, T., 2019. Challenges with availability and quality of climate data in Africa. Elsevier eBooks, pp. 71–80. https://doi.org/10.1016/b978-0-12-815998-9.00007-5.

Du, H., Tan, M.L., Zhang, F., Chun, K.P., Li, L., Kabir, M.H., 2023. Evaluating the effectiveness of CHIRPS data for hydroclimatic studies. Theor. Appl. Climatol. 155 (3), 1519–1539. https://doi.org/10.1007/s00704-023-04721-9.

Dunning, C.M., Black, E.C.L., Allan, R.P., 2016. The onset and cessation of seasonal rainfall over Africa. J. Geophys. Res.: Atmos. 121 (19). https://doi.org/10.1002/2016jd025428. https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2016JD025428.

Ensor, L.A., Robeson, S.M., 2008. Statistical characteristics of daily precipitation: comparisons of gridded and point datasets. J. Appl. Meteorol. Climatol. 47 (9), 2468–2476. https://doi.org/10.1175/2008jamc1757.1.

Fischer, G., Nachtergaele, F., H, V.V., Chiozza, F., Francheschini, G., Henry, M., Muchoney, D., Tramberend, S., 2021. Global agro-ecological zone V4 – Model documentation. FAO eBooks. https://doi.org/10.4060/cb4744en.

Frieler, K., Schauberger, B., Arneth, A., Balkovič, J., Chryssanthacopoulos, J., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Müller, C., Olin, S., Pugh, T. a M., Schaphoff, S., Schewe, J., Schmid, E., Warszawski, L., Levermann, A., 2017. Understanding the weather signal in national crop-yield variability. Earth's Future 5 (6), 605–616. https://doi.org/10.1002/2016ef000525.

Gocic, M., Trajkovic, S., 2013. Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia. Glob. Planet. Change 100, 172–182. https://doi.org/10.1016/j.gloplacha.2012.10.014. https://linkinghub.elsevier.com/retrieve/pii/S0921818112002032.

Gummadi, S., Dinku, T., Shirsath, P.B., Kadiyala, M.D.M., 2022. Evaluation of multiple satellite precipitation products for rainfed maize production systems over Vietnam. Sci. Rep. 12 (1), 485. https://doi.org/10.1038/s41598-021-04380-8.

Gwambene, B., Liwenga, E.T., Mung'ong'o, C.G., 2022. Climate Change and Variability Impacts on Agricultural Production and Food Security for the Smallholder Farmers in Rungwe, Tanzania. Environ. Manag. 71 (1), 3–14. https://doi.org/10.1007/s00267-022-01628-5.

Hoffmann, M., Haakana, M., Asseng, S., Höhn, J., Palosuo, T., Ruiz-Ramos, M., Fronzek, S., Ewert, F., Gaiser, T., Kassie, B., Paff, K., Rezaei, E., Rodríguez, A., Semenov, M., Srivastava, A., Stratonovitch, P., Tao, F., Chen, Y., Rötter, R., 2017. How does inter-annual variability of attainable yield affect the magnitude of yield gaps for wheat and maize? An analysis at ten sites. Agric. Syst. 159, 199–208. https://doi.org/10.1016/j.agsy.2017.03.012.

IFDC. (2025, January). AfricaFertilizer. Retrieved February 7, 2025, from https://africafertilizer.org/#/en.

Joseph, J.E., Akinseye, F.M., Worou, O.N., Faye, A., Konte, O., Whitbread, A.M., Rötter, R.P., 2023. Assessment of the relations between crop yield variability and the onset and intensity of the West African Monsoon. Agric. For. Meteorol. 333, 109431. https://doi.org/10.1016/j.agrformet.2023.109431.

Josephson, A., Michler, J.D., Kilic, T., Murray, S., 2026. The mismeasure of weather: Using earth observation data for estimation of socioeconomic outcomes. J. Dev. Econ. 178, 103553. https://doi.org/10.1016/j.jdeveco.2025.103553.

Keating, B., Carberry, P., Hammer, G., Probert, M., Robertson, M., Holzworth, D., Huth, N., Hargreaves, J., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J., Silburn, M., Wang, E., Brown, S., Bristow, K., Asseng, S., Smith, C., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18 (3–4), 267–288. https://doi.org/10.1016/s1161-0301(02)00108-9.

Kessabi, R., Hanchane, M., Caloiero, T., Pellicone, G., Addou, R., Krakauer, N.Y., 2023. Analyzing spatial trends of precipitation using gridded data in the Fez-Meknes region, Morocco. Hydrology 10 (2), 37. https://doi.org/10.3390/hydrology10020037.

MacCarthy, D.S., Adam, M., Freduah, B.S., Fosu-Mensah, B.Y., Ampim, P. a Y., Ly, M., Traoré, P.C.S., Adiku, S.G., 2021. Climate change impact and variability on cereal productivity among smallholder farmers under future production systems in West Africa. Sustainability 13 (9), 5191. https://doi.org/10.3390/su13095191.

Mankin, K.R., Mehan, S., Green, T.R., Barnard, D.M., 2025. Review of gridded climate products and their use in hydrological analyses reveals overlaps, gaps, and the need for a more objective approach to selecting model forcing datasets. Hydrol. Earth Syst. Sci. 29 (1), 85–108. https://doi.org/10.5194/hess-29-85-2025.

Mathevet, T., Moine, N.L., Andréassian, V., Gupta, H., Oudin, L., 2023. Multi-objective assessment of hydrological model performances using Nash–Sutcliffe and Kling–Gupta efficiencies on a worldwide large sample of watersheds. Comptes Rendus Géoscience 355 (S1), 117–141. https://doi.org/10.5802/crgeos.189.

Mavromatis, T., 2008. Estimation of solar radiation and its application to crop simulation models in Greece. Clim. Res. 36, 219–230. https://doi.org/10.3354/cr00750. http://www.int-res.com/abstracts/cr/v36/n3/p219-230/.

Ministry of Agriculture. (2025, January 7). Ministry of Agriculture, United Republic of Tanzania. Retrieved January 7, 2025, from https://www.kilimo.go.tz.

Mourtzinis, S., Edreira, J.I.R., Conley, S.P., Grassini, P., 2016. From grid to field: Assessing quality of gridded weather data for agricultural applications. Eur. J. Agron. 82, 163–172. https://doi.org/10.1016/j.eja.2016.10.013.

Msongaleli, B.M., Rwehumbiza, F., Tumbo, S.D., Kihupi, N., 2015. Impacts of climate variability and change on rainfed sorghum and maize: Implications for food security policy in Tanzania. J. Agric. Sci. 7 (5). https://doi.org/10.5539/jas.v7n5p124.

Nassaj, B.N., Zohrabi, N., Shahbazi, A.N., Fathian, H., 2022. Evaluating the performance of eight global gridded precipitation datasets across Iran. Dyn. Atmospheres Oceans 98, 101297. https://doi.org/10.1016/j.dynatmoce.2022.101297.

NCAR. (2022, September 9). CPC Unified Gauge-Based Analysis of Global Daily Precipitation | Climate Data Guide. https://climatedataguide.ucar.edu/climate-data/cpc-unified-gauge-based-analysis-global-daily-precipitation.

Ogunbode, T.O., Esan, V.I., Samson, T.K., Oyelowo, O.J., Asifat, J.T., 2022. Rainfall trend and its implications for sustainable crop production and water resources management: a case study of Iwo, Nigeria. J. Appl. Sci. Environ. Manag. 26 (8), 1415–1422. https://doi.org/10.4314/jasem.v26i8.15.

Reichle, R.H., Liu, Q., Koster, R.D., Draper, C.S., Mahanama, S.P.P., Partyka, G.S., 2017. Land surface precipitation in MERRA-2. J. Clim. 30 (5), 1643–1664. https://doi.org/10.1175/jcli-d-16-0570.1.

Riyanto, S., Sitanggang, I.S., Djatna, T., Atikah, T.D., 2023. Comparative analysis using various performance metrics in imbalanced data for multi-class text classification. Int. J. Adv. Comput. Sci. Appl. 14 (6). https://doi.org/10.14569/ijacsa.2023.01406116.

Rötter, R., Van Keulen, H., 1997. Variations in yield response to fertilizer application in the tropics: II. Risks and opportunities for smallholders cultivating maize on Kenya's arable land. Agric. Syst. 53 (1), 69–95. https://doi.org/10.1016/s0308-521x(96)00037-6.

Satgé, F., Defrance, D., Sultan, B., Bonnet, M., Seyler, F., Rouche, N., Pierron, F., Paturel, J., 2020. Evaluation of 23 gridded precipitation datasets across West Africa. J. Hydrol. 581, 124412. https://doi.org/10.1016/j.jhydrol.2019.124412.

Saxton, K.E., Rawls, W.J., 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil. Sci. Soc. Am. J. 70 (5), 1569–1578. https://doi.org/10.2136/sssaj2005.0117.

Schwalbert, R., Amado, T.J., Horbe, T. a N., Stefanello, L.O., Assefa, Y., Prasad, P., Rice, C.W., Ciampitti, I.A., 2018. Corn yield response to plant density and nitrogen: spatial models and yield distribution. Agron. J. 110 (3), 970–982. https://doi.org/10.2134/agronj2017.07.0425.

Selina Wamucii, 2025. Glob. 2023 Prices. Trends & Insights|Agric. Stat. https://www.selinawamucii.com/insights/prices/senegal January 23.

Sheahan, M., Barrett, C.B., 2016. Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Policy 67, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010.

Stern, R.D., Cooper, P.J.M., 2011. Assessing climate risk and climate change using rainfall data – a case study from Zambia. Exp. Agric. 47 (2), 241–266. https://doi.org/10.1017/s0014479711000081.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107. https://doi.org/10.1002/2017rg000574.

Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R.P., Stringer, M., Chadwick, R., Kayitakire, F., 2014. Extension of the TAMSAT Satellite-Based Rainfall Monitoring over Africa and from 1983 to Present. J. Appl. Meteorol. Climatol. 53 (12), 2805–2822. https://doi.org/10.1175/jamc-d-14-0016.1.

Tittonell, P., Giller, K.E., 2012. When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. Field Crops Res. 143, 76–90. https://doi.org/10.1016/j.fcr.2012.10.007.

Tonnang, H.E., Hervé, B.D., Biber-Freudenberger, L., Salifu, D., Subramanian, S., Ngowi, V.B., Guimapi, R.Y., Anani, B., Kakmeni, F.M., Affognon, H., Niassy, S., Landmann, T., Ndjomatchoua, F.T., Pedro, S.A., Johansson, T., Tanga, C.M., Nana, P., Fiaboe, K.M., Mohamed, S.F., Borgemeister, C., 2017. Advances in crop insect modelling methods—Towards a whole system approach. Ecol. Model. 354, 88–103. https://doi.org/10.1016/j.ecolmodel.2017.03.015.

Wang, W., Hocke, K., Mätzler, C., 2021. Physical Retrieval of Rain Rate from Ground-Based Microwave Radiometry. Remote. Sens. 13 (11), 2217. https://doi.org/10.3390/rs13112217.

WMO (2020). Retrieved January 7, 2025, from https://community.wmo.int/en/activity-areas/climate-services/climate-products-and-initiatives/wmo-climatological-normals.

World Meteorological Organization. (2017). *WMO Guidelines on the Calculation of Climate Normals: Weather Climate Water* (2017 edition, Vol. 1203). https://library.wmo.int/doc_num.php?explnum_id=4166.

Zandler, H., Haag, I., Samimi, C., 2019. Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. Sci. Rep. 9 (1). https://doi.org/10.1038/s41598-019-51666-z.

Zhang, G., Shen, D., Xie, R., Ming, B., Hou, P., Xue, J., Li, R., Chen, J., Wang, K., Li, S., 2020. Optimizing planting density to improve nitrogen use of super high-yield maize. Agron. J. 112 (5), 4147–4158. https://doi.org/10.1002/agj2.20334.