

Machine Learning-Based Yield and Water Efficiency Prediction for Deficit-Irrigated Maize in Semi-Arid Ethiopia

Melaku Adugnaw Walle¹

melakuadugnaw2010@gmail.com

Arba Minch University

Hailemariam Molla Ashagre²

Arba Minch University

Research Article

Keywords: Deficit irrigation, Irrigation management, Machine learning, Maize, Water use efficiency, Yield prediction

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11242111/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Machine Learning-Based Yield and Water Efficiency Prediction for Deficit-Irrigated Maize in Semi-Arid Ethiopia

Melaku Adugnaw Walle^{1*} and Hailemariam Molla Ashagre²

¹ Faculty of Water Resources and Irrigation Engineering, Arba Minch Water Technology Institute, Arba Minch University, Arba Minch, Ethiopia

² Faculty of Hydraulic and Water Resources Engineering, Arba Minch Water Technology Institute, Arba Minch University, Arba Minch, Ethiopia

Corresponding author*: melakuadugnaw2010@gmail.com

Highlights

- Deficit irrigation: Applying 70% ET_c produced 6,621 kg ha⁻¹ grain yield while saving 1,350 m³ ha⁻¹ (30%) of irrigation water compared with full irrigation.
- Grain-yield prediction: XGBoost showed the highest test-set performance for maize grain-yield prediction ($R^2 = 0.8912$, RMSE = 183.01 kg ha⁻¹, MAE = 155.05 kg ha⁻¹) using temporal crop-growth features.
- WUE prediction: Ridge Regression showed the highest test-set performance for water use efficiency prediction ($R^2 = 0.6540$, RMSE = 0.210 kg m⁻³, MAE = 0.171 kg m⁻³).
- Key predictors: Feature-importance analysis identified biomass growth rate (47.7%) as the most important predictor of grain yield and cob shape ratio (43.6%) as the most important predictor of WUE.

Abstract

Water scarcity constrains maize production in semi-arid Ethiopia, requiring irrigation strategies that improve water productivity while maintaining grain yield. This study integrated machine learning with water-balance-based deficit irrigation to predict maize grain yield and water use efficiency (WUE). Field experiments evaluated five irrigation treatments (100%, 85%, 70%, 55%, and 40% ET_c) in a randomized complete block design. Plant height, leaf length, and biomass were measured as growth parameters, while cob length, cob diameter, number of grains per cob, and grain yield were evaluated as yield-related parameters, together with irrigation water applied. Random Forest, XGBoost, Linear Regression, and Ridge Regression models were evaluated. XGBoost provided the highest test-set accuracy for grain yield prediction ($R^2 = 0.8912$, RMSE = 183.01 kg ha⁻¹, MAE = 155.05 kg ha⁻¹), whereas Ridge Regression performed best for WUE prediction ($R^2 = 0.6540$, RMSE = 0.210 kg m⁻³, MAE = 0.171 kg m⁻³). Under the

experimental conditions, 70% ET_c provided a balance between grain yield and water saving, producing 6,621 kg ha⁻¹ with a WUE of 2.10 kg m⁻³ and saving 1,350 m³ ha⁻¹ (30%) relative to full irrigation. Given the small dataset, the ML results are exploratory and require validation using larger multi-season and multi-location datasets.

Keywords: Deficit irrigation; Irrigation management; Machine learning; Maize; Water use efficiency; Yield prediction

1. Introduction

Water scarcity is an escalating global challenge that increasingly constrains agricultural production, particularly in semi-arid regions where rainfall is limited, highly variable, and strongly influenced by climate change (McDonald et al. 2011; El-Rawy et al. 2023)). Rising temperatures, erratic precipitation patterns, and increasing frequency of drought events are intensifying pressure on freshwater resources and accelerating soil moisture depletion, thereby undermining agricultural productivity and long-term system stability (Intergovernmental Panel On Climate Change (IPCC), 2023); Muharomah et al., 2025). Although agriculture remains the largest consumer of freshwater, irrigation efficiency in many developing regions remains low, resulting in substantial water losses, declining productivity, and increasing competition for limited resources (Morante-Carballo et al., 2022; Ben Hassen et al., 2025). In Ethiopia, these challenges are particularly severe. Agriculture dominates water use and remains central to livelihoods and the national economy, yet irrigation systems are often inefficient due to infrastructural constraints, weak management practices, and continued reliance on conventional scheduling approaches that do not reflect actual crop water demand (Birhanu et al. 2021; Gebeyhu Reta et al. 2024). In the areas surrounding Arba Minch, water availability is increasingly constrained by climate variability, declining rainfall, and seasonal reductions in river flow, which together are intensifying pressure on water resources and reducing supply reliability (Yemane et al., 2015; Setu et al., 2023). Consequently, water scarcity has become not only a resource limitation but also a major driver of production risk, yield instability, and economic uncertainty.

Maize (*Zea mays* L.) is a dominant staple crop in this system and a critical component of food security. However, its productivity remains substantially below potential levels Assefa et al.(2020) This yield gap is strongly linked to inefficient water management and the crop's high sensitivity to water stress during critical growth stages, including tasseling, silking, and grain filling ((Tesfayeteferayigezu et al. 2016; Chinasho et al. 2023; Pandey et al. 2000; Oktem, 2008; Karasu et al. 2015). Water deficits during these stages disrupt physiological processes such as pollination, assimilate partitioning, and kernel development, leading to reduced grain formation and significant yield losses (Farré & Faci, 2009). Furthermore, increasing temperature extremes and rainfall variability are exacerbating these effects, contributing to higher yield variability and

long-term productivity decline in maize-based systems (Lobell et al., 2011; Intergovernmental Panel On Climate Change (IPCC), 2023)

Under such conditions, improving water productivity while sustaining yield has become a critical research and policy priority. Deficit irrigation, when properly aligned with crop water requirements, has been widely recognized as an effective strategy for optimizing water use under limited water availability (Feres & Soriano, 2006). Its performance is further enhanced when integrated with water-balance-based scheduling approaches that account for soil moisture dynamics, evapotranspiration demand, and rainfall inputs to guide irrigation timing and application depth (Pereira et al. 2015; Pragg et al. 2024). Empirical evidence indicates that such integrated approaches can improve water use efficiency and reduce non-productive water losses while maintaining acceptable yield levels when carefully managed (Igbadun et al. 2008; Zhang et al. 2019; Zou et al. 2021). However, their practical adoption remains limited in data-scarce environments, largely due to the absence of locally validated decision-support tools and predictive frameworks.

In parallel, advances in machine learning have created new opportunities for improving agricultural water management and yield prediction. Machine learning techniques are capable of capturing complex, non-linear interactions between crop growth dynamics, environmental conditions, and management practices, which are often not adequately represented in conventional statistical approaches (Chlingaryan et al. 2018; Shahhosseini et al. 2021; Van Klompenburg et al. 2020). Temporal growth variables plant height, leaf length, and biomass provide dynamic information on crop development that can improve the prediction of maize yield and water use efficiency. Similarly, yield attributes such as cob length, cob diameter, and number of grain per cob are important indicators of maize productivity and are highly responsive to water availability (Melkie et al. 2024; Zhang et al. 2019; Jeong et al. 2016). Currently, few studies have developed predictive frameworks specifically designed to estimate maize grain yield and water use efficiency (WUE) from temporal crop-growth measurements under deficit irrigation in Ethiopia. This represents an important methodological gap because most existing machine-learning applications for maize yield prediction have emphasized remote-sensing, environmental, or historical yield variables, rather than integrating repeated field-based crop-growth measurements with irrigation-management information. Recent systematic evidence further indicates substantial geographical and methodological disparities in the application of

machine learning to maize yield prediction, with research concentrated in a limited number of countries and relatively fewer studies from Africa Agyeman et al. (2025) In Ethiopia, recent machine-learning research has primarily emphasized remote-sensing-based maize yield estimation. Mondschein et al. (2024) applied Sentinel-2 imagery and vegetation indices to estimate field-level maize yield in Ethiopian smallholder farming systems, with the best-performing models achieving an (R^2) of 0.63. However, model performance varied across regions, indicating limitations in spatial transferability under varying environmental and image-quality conditions. Moreover, the study focused on remote-sensing-based yield estimation and did not incorporate deficit-irrigation treatments or predict irrigation water use efficiency, leaving these aspects insufficiently addressed in the Ethiopian context. More broadly, reviews of artificial-intelligence applications in agricultural water management indicate that machine-learning approaches have been increasingly explored for irrigation and water-management decisions, but their applicability remains constrained by data availability, contextual variability, and limited validation across diverse production environments Tace et al. (2022). These limitations are particularly relevant to data-constrained agricultural systems in sub-Saharan Africa, where locally collected field data and management-specific observations are essential for developing contextually reliable predictive models. Therefore, a specific research gap remains in the development and comparative evaluation of machine-learning models that use temporal crop-growth characteristics together with deficit-irrigation treatments to predict both maize grain yield and WUE under Ethiopian field conditions. Addressing this gap could provide a more management-oriented predictive framework for evaluating alternative irrigation regimes and identifying water-saving strategies while maintaining acceptable maize productivity.

This study integrates machine learning with water-balance-based deficit irrigation to predict maize grain yield and water use efficiency (WUE) under semi-arid conditions in Ethiopia. It combines irrigation-management treatments with temporally resolved crop-growth data and compares multiple machine-learning algorithms using field-experiment observations. The study provides an exploratory assessment of crop-growth-based machine learning for irrigation-management analysis under data-limited conditions. The objectives were to: (1) evaluate the effects of irrigation levels on maize growth, grain yield, and WUE; (2) develop and compare machine-learning models using temporal crop-growth data; and (3) identify an irrigation level that balances water savings and grain yield.

2. Materials and Methods

2.1. Description of the Study Area

The study was conducted at Arba Minch university demonstration farm, located in Arba Minch Zuria Woreda, Gamo Zone, South Ethiopia Regional State, Ethiopia. Geographically, the study area is situated between 5°50' and 6°30' N latitude and 37°20' and 38°00' E longitude, as indicated by the geographical extent shown in Fig 1.

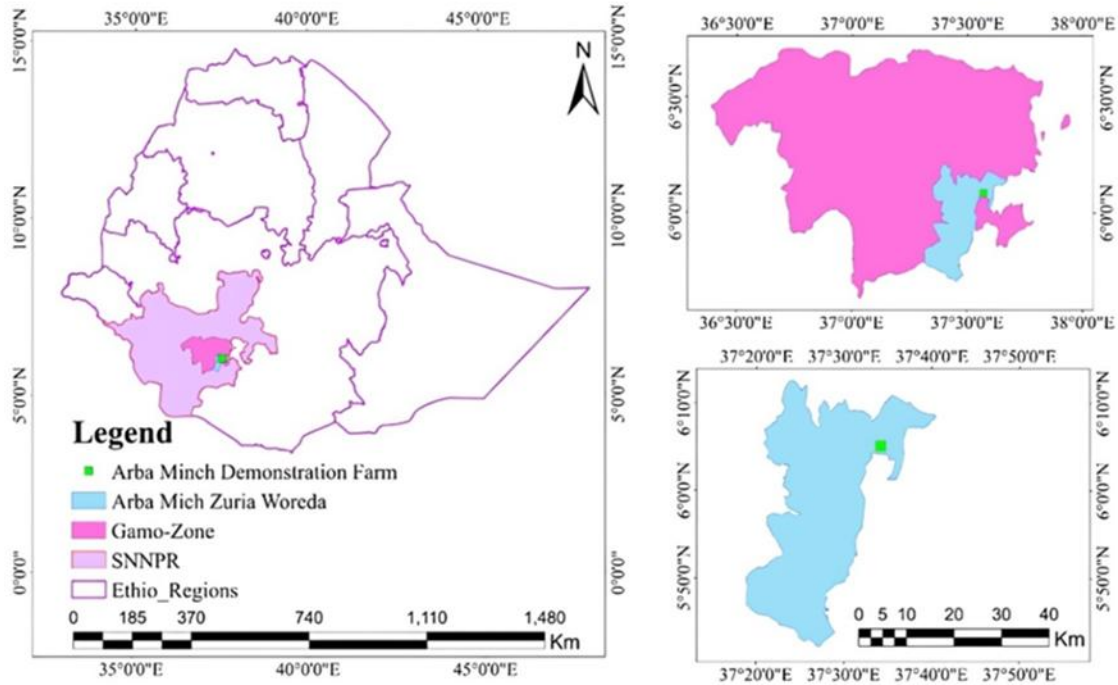


Fig. 1 Location of Arba Minch Demonstration Farm, South Ethiopia Regional State, Ethiopia

2.2. Climatic Characteristics of the Study Area

2.2.1. Long-Term Climatic Patterns

Long-term monthly meteorological data, including rainfall, maximum temperature (Tmax), minimum temperature (Tmin), and average temperature (Tavg), were obtained from the Arba Minch University Meteorological Station for the period 1987–2018. The 32-year climatic record revealed considerable seasonal variation in rainfall and temperature. Monthly rainfall ranged from approximately 23 to 163 mm, while Tmax, Tmin, and Tavg ranged from 27.5 to 34 °C, 15 to 19 °C, and 23 to 26 °C, respectively (Fig. 2). These variations highlight the seasonal climatic characteristics of the study area and provide a long-term climatic context for the experiment.

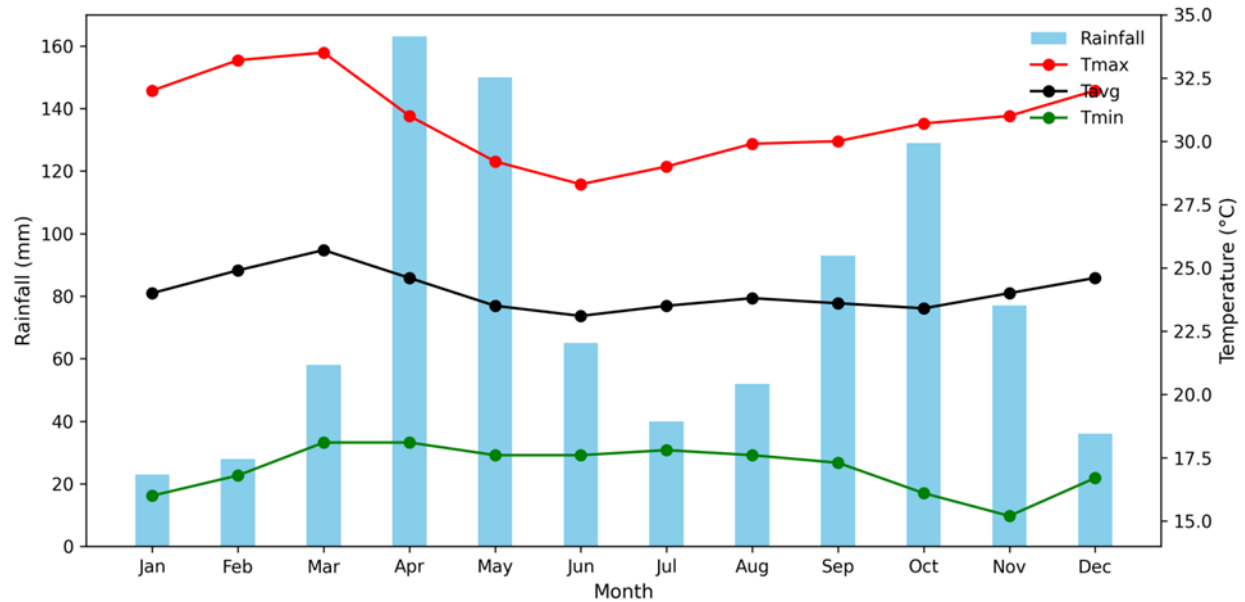


Fig. 2 Long-term monthly rainfall, maximum temperature (Tmax), minimum temperature (Tmin), and average temperature (Tavg) patterns in the study area based on meteorological records from 1987 to 2018

2.2.2. Meteorological Conditions during the Experimental Period

Meteorological conditions during the experimental period (November–February) showed variations in temperature, wind speed, relative humidity, solar radiation, and sunshine duration (Table 1). Maximum and minimum temperatures increased progressively from November to February, ranging from 29.0 to 33.0 °C and 16.4 to 20.0 °C, respectively. Similarly, solar radiation and sunshine duration increased from 6.7 to 9.0 MJ m⁻² day⁻¹ and 6.8 to 9.4 hours per day, respectively. Relative humidity fluctuated between 52.1% and 68.7%, whereas wind speed ranged from 0.20 to 0.60 m s⁻¹. Overall, these meteorological variations characterize the environmental conditions prevailing during the crop-growing period and provide context for interpreting crop growth, grain yield, and water-use efficiency.

Table 1 Monthly meteorological condition during the experimental period (November–February).

Months	Temperature (°C)		Wind speed (m s ⁻¹)	Relative humidity (%)	Solar Radiation	Sun shine hour (h)
	Tmax	Tmin			(MJ m ⁻² day ⁻¹)	
November	29	16.4	0.45	61.8	6.7	6.8
December	31	17.6	0.6	65	7.5	7.5
January	31.3	18.2	0.2	52.1	8.5	8.6

2.3. Soil Characteristics

Composite soil samples were collected from four depth intervals (0–30, 30–60, 60–90, and 90–120 cm) and analyzed for physical and chemical properties, including particle-size distribution, bulk density, electrical conductivity (EC), pH, organic matter, field capacity (FC), and permanent wilting point (PWP). The soil profile was predominantly clay-textured, with bulk density ranging from 0.93 to 1.12 g cm⁻³. Soil pH ranged from 8.1 to 8.3, while EC remained low (0.058–0.060 dS m⁻¹). Field capacity and permanent wilting point ranged from 31.4–34.7% and 16.8–21.3% (vol.), respectively, indicating relatively high water-retention capacity (Table 2).

Table 2 Soil physical properties of the experimental plots at different depths.

Soil depth (cm)	Bulk density (g cm ⁻³)	% clay	% silt	% sand	Texture	EC (dS/m)	pH	Organic Matter (%)	FC (Vol. %)	PWP (Vol. %)
0–30	0.93	51.5	39	9.5	Clay	0.059	8.3	13	34.2	21.3
30–60	1.12	48.5	37	14.5	Clay	0.059	8.1	13.9	34.7	19.7
60–90	1.08	51.6	40	8.4	Clay	0.058	8.1	12.6	31.6	17
90–120	1.1	59.3	20	20.7	Clay	0.06	8.1	12.4	31.4	16.8
Average	1.06	52.73	34.00	13.28	Clay	0.06	8.15	12.98	32.98	18.70

FC = Field capacity; PWP = Permanent wilting point; TAW = Total available water.

2.4. Experimental Design and Treatments

Field experiments were conducted using a Randomized Complete Block Design (RCBD) over a total area of 31.0 m × 28.0 m. The experiment consisted of three replication blocks (R1, R2, and R3), each containing five randomly assigned irrigation treatments (T1–T5) to account for spatial variability. Each experimental plot measured 5.0 m × 8.0 m (40 m²), with 1.0 m buffer zones between adjacent plots to minimize lateral water movement and root interference. Maize was planted on raised ridges, and irrigation water was supplied from a central source through manifold feeder lines and distributed to individual plot furrows for treatment-specific application. The experimental layout and arrangement of irrigation treatments are shown in Fig. 3.

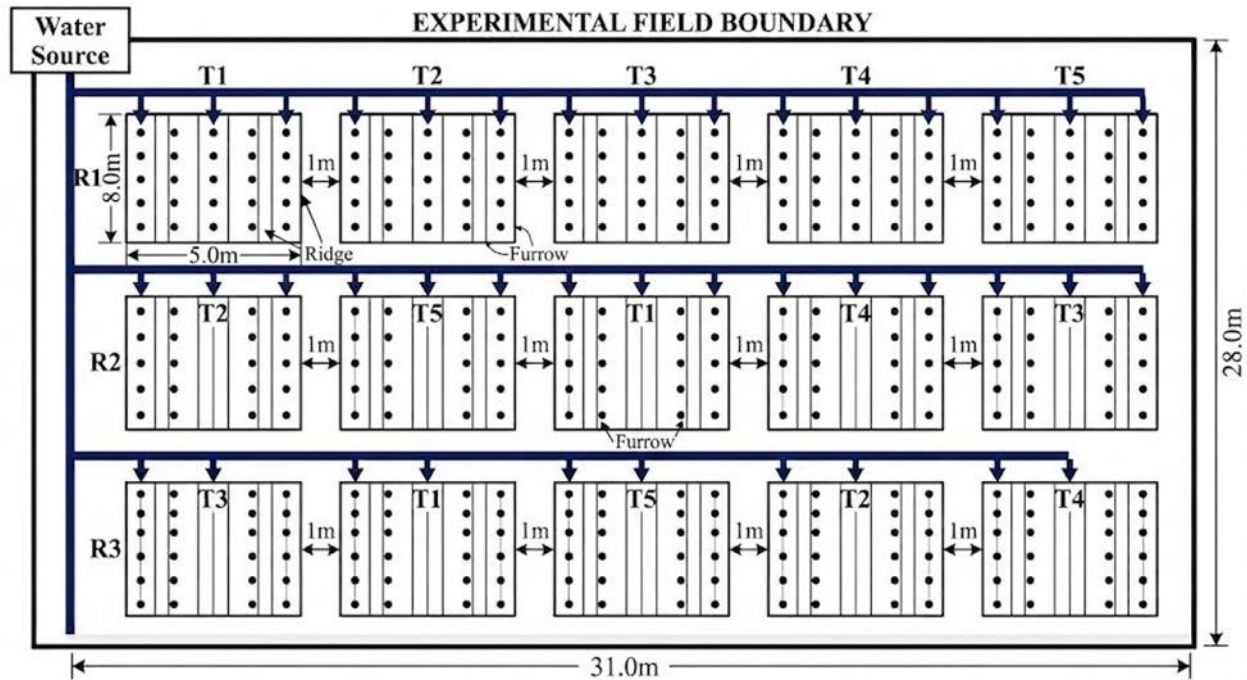


Fig. 3 Experimental field layout showing the randomized irrigation treatments and replication blocks

The irrigation treatments were established as percentages of crop evapotranspiration (ET_c), with T1 representing the full irrigation treatment (100% ET_c) and T2–T5 represents deficit irrigation levels of 85%, 70%, 55%, and 40% ET_c, respectively (Table 3).

Table 3 Irrigation treatments and corresponding percentage of crop evapotranspiration (ET_c).

Treatments	% ET _c
T1	100 (control treatment)
T2	85
T3	70
T4	55
T5	40

2.5. Irrigation Scheduling

Irrigation scheduling was based on the daily soil water balance method, which accounts for precipitation, irrigation input, crop evapotranspiration, and the resulting change in soil water storage (Lee & Jung, 2018; Zhao et al. 2022) as expressed in Equation 1.

$$\Delta S = P + I - ET - R - D \dots \dots \dots 1$$

Where:- ΔS represents the change in soil water storage (m), P is precipitation (m), I is irrigation water applied (m), ET is crop evapotranspiration (m), R is runoff (m), and D is deep percolation

(m). Runoff was assumed negligible because of the low field slope and blocked-end furrow configuration. Because irrigation was scheduled to replenish soil water only to the predetermined allowable depletion level, deep percolation was assumed negligible and set to zero in the daily water-balance calculation.

2.5.1. Reference Evapotranspiration (ET_o)

Reference evapotranspiration (ET_o) was estimated using the Penman–Monteith method, which combines temperature, solar radiation, humidity, and wind speed to provide an accurate measure of potential water loss from a well-watered reference surface, as expressed in Equation (2).

$$ET_o = \frac{0.408\Delta(Rn - G) + \frac{\gamma * 900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \dots\dots\dots 2$$

Where: - The reference crop evapotranspiration (ET_o) was calculated in mm.day⁻¹. In this study, Δ represents the slope of the saturation vapor pressure curve (kPa.°C⁻¹), R_n is the net radiation at the crop surface (MJ.m⁻².day⁻¹), and G is the soil heat flux density (MJ.m⁻².day⁻¹). T is the mean daily air temperature at 2 m height (°C), U₂ is the wind speed at 2 m height (m.s⁻¹), and e_s–e_a is the saturation vapor pressure deficit (kPa), where e_s is the saturation vapor pressure at a given period (kPa) and e_a is the actual vapor pressure (kPa), γ denotes the psychrometric constant (kPa.°C⁻¹).

2.5.2. Crop Evapotranspiration

Crop evapotranspiration (ET_c) was estimated using the crop coefficient approach (Equation 3):

$$ET_c = k_c * ET_o \dots\dots\dots 3$$

Where K_c represents the crop coefficient of maize, and ET_o denotes the reference evapotranspiration (mm day⁻¹).

Crop coefficient (K_c) values were assigned to the initial, development, mid-season, and late growth stages based on published maize-specific crop coefficient estimates. Abedinpour, (2015) reported growth-stage-specific K_c values of 0.40–0.60, 0.70–0.80, 1.10–1.21, and 0.50–0.65 for the initial, development, mid-season, and late-season stages, respectively. In the present study, K_c values of 0.3, 0.5, 1.2, and 0.5 were adopted for these stages, respectively, consistent with commonly used maize crop-coefficient values. The timing and duration of each growth stage were determined from crop establishment dates and field observations.

2.6. Irrigation Water Application

Irrigation water was diverted from the Kulfo River, conveyed through lined canals, and applied to the experimental plots using furrow irrigation. For each treatment, the target ET_c was calculated as 1.00, 0.85, 0.70, 0.55, or 0.40 of the estimated ET_c for the 100%, 85%, 70%, 55%, and 40% ET_c treatments, respectively. Irrigation requirements were determined using the soil-water balance, with the 100% ET_c treatment serving as the reference. The calculated irrigation depth was subsequently adjusted according to the assigned treatment fraction and converted to the corresponding irrigation volume for each plot.

The irrigation volume applied to each plot was calculated using (Equation 4):

$$V = \frac{I_d * A}{1000} \dots\dots\dots 4$$

Where V represents the irrigation volume (m³), I_d irrigation depth (mm) and A the plot area (m²).

2.7. Agronomic Practices and Measurements

Growth parameters were monitored at 15-day intervals from crop establishment through physiological maturity. In each experimental plot, measurements were recorded from ten randomly selected representative plants to ensure a reliable estimate of within-plot variability. Yield-related traits, including cob length, cob diameter, number of grains per cob, and grain yield, were determined following harvest.

2.8. Water Use Efficiency

Water use efficiency (WUE) was calculated as grain yield per unit of irrigation water applied, as expressed in Equation (5) Howell, (2001).

$$WUE = \frac{Y}{I} \dots\dots\dots 5$$

Where, WUE is irrigation water use efficiency (kg m⁻³), Y is grain yield (kg ha⁻¹), and I is applied irrigation water (m³ ha⁻¹).

2.9. Machine Learning Algorithms, Feature Engineering, and Feature Importance

The machine learning analysis was conducted in Python (v3.13.15) using Google Colab, with scikit-learn (v1.6.1), XGBoost (v3.4.1), pandas (v2.2.3), NumPy (v2.1.3), and Matplotlib (v3.10.0). Four algorithms—Random Forest (RF), XGBoost (XGB), Linear Regression (LR), and Ridge Regression (RR)—were used to predict maize grain yield and water use efficiency

(WUE). RF and XGB were used to capture nonlinear relationships, whereas LR and RR provided linear baselines, with RR incorporating L2 regularization to reduce over-fitting.

The experiment comprised 15 plots representing five irrigation treatments with three replications. Plant height, leaf length, and biomass were measured at eight sampling dates (15, 30, 45, 60, 75, 90, 105, and 120 days after planting), resulting in 120 temporal records. Raw field measurements were transformed into a plot-level predictor matrix through a structured feature-engineering process. Static features included final, maximum, and mean values and cob shape ratio, whereas dynamic features included growth rate, early- and late-season growth, and area under the curve (AUC). AUC was calculated using the trapezoidal method Ntouros et al. (2025). Growth rate was calculated from the change between initial and final measurements relative to the corresponding time interval Colusse et al. (2020). The feature-engineering process is illustrated in Fig. 4.

$$AUC = \sum_{i=1}^{n-1} \frac{X_i + X_{i+1}}{2} (t_{i+1} - t_i) \dots \dots \dots 6$$

Growth rate was calculated as:

$$GR = \frac{X_f - X_i}{t_f - t_i} \dots \dots \dots 7$$

The mean and maximum values were calculated as:

$$X_{mean} = \frac{1}{n} \sum_{i=1}^n X_i \dots \dots \dots 8$$

$$X_{max} = \text{Max}(X_1, X_2, \dots, X_n) \dots \dots \dots 9$$

Where X is the measured crop-growth variable, X_i and X_f are the initial and final measurements, respectively, t_i and t_f are their corresponding measurement times, X_1 – X_n are the sequential measurements, n is the total number of measurements, and X_{max} is the maximum observed value.

End-of-season variables, including cob length, cob diameter, number of grains per cob, harvest index, and total water applied, were initially considered as candidate predictors. Number of grains per cob and harvest index was excluded from grain-yield prediction to minimize data leakage because they were measured at or near harvest and are directly related to final grain yield. Total water applied was excluded from WUE prediction to avoid direct arithmetic circularity because WUE was calculated from grain yield and applied irrigation water. Separate

predictor sets were therefore developed for grain yield and WUE. The resulting feature matrix was divided into 70% training and 30% independent testing subsets, with five-fold cross-validation applied during model development.

Feature importance analysis was conducted to identify predictors contributing to model predictions. For RF, importance was determined using mean decrease in impurity, whereas XGB importance was assessed based on gain. Predictors were ranked according to their relative contribution to grain-yield and WUE predictions.

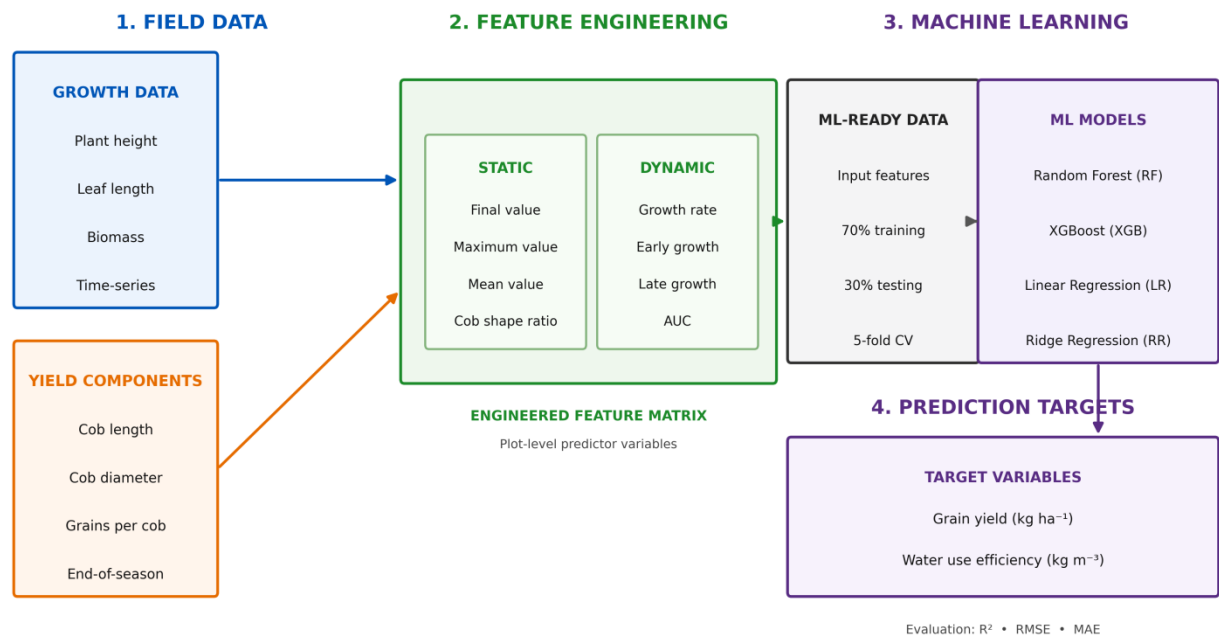


Fig. 4 Workflow of feature engineering and dataset construction for machine learning prediction of maize yield and water use efficiency.

2.10. Cross-Validation, Feature Scaling, and Model Testing

The dataset was divided into 70% training ($n = 10$) and 30% independent testing ($n = 5$) subsets using a fixed random seed to ensure reproducibility. The training subset was used for model development and five-fold cross-validation, while the testing subset was withheld until final model evaluation. During cross-validation, the training data were divided into five folds, with each fold used once for validation and the remaining four folds used for model training. Final models were fitted using the complete training dataset and evaluated on the independent testing dataset Fig. 5.

Feature scaling using z-score standardization was applied only to the Linear Regression and Ridge Regression models. Random Forest and XGBoost are tree-based methods and therefore did not require scaling. To prevent data leakage, scaling parameters were estimated exclusively from the corresponding training data. During five-fold cross-validation, the scaler was fitted separately within each training fold and then applied to its corresponding validation fold. After cross-validation, the scaler was refitted using the complete training dataset and applied to both the training and independent testing data.

Given the small sample size of 15 experimental plots, the resulting training and testing subsets were limited. Therefore, the machine learning analysis was considered exploratory, providing evidence of model behavior under the experimental conditions rather than establishing broad predictive generalizability. Validation using larger datasets covering multiple seasons, locations, and experimental conditions would be required for wider application.

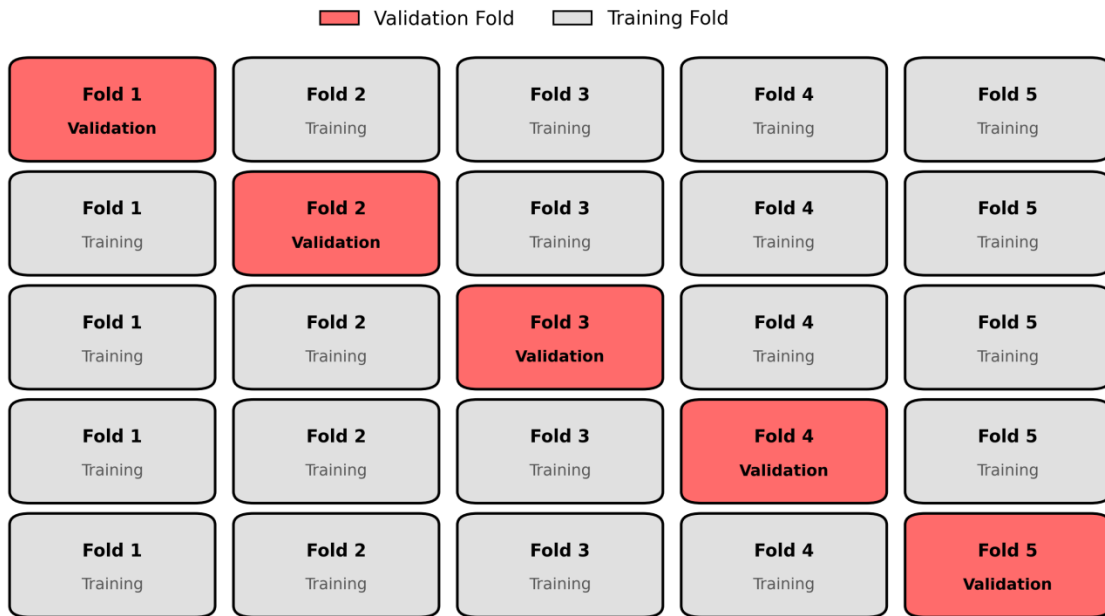


Fig. 5 Five-fold cross-validation procedure

2.11. Prediction Targets and Model Evaluation

Two prediction tasks were developed: maize grain yield and water use efficiency (WUE). For grain-yield prediction, the final predictor set consisted of crop growth and yield-related features that were available without directly using the target variable. Number of grains per cob and harvest index were excluded because of their close association with final grain yield and their potential to introduce data leakage. For WUE prediction, total water applied was additionally

excluded because WUE was calculated from grain yield and irrigation water applied, thereby avoiding direct arithmetic circularity. Thus, WUE was predicted using the remaining growth and yield-related features.

The predictive performance of RF, XGB, LR, and RR was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). These metrics were calculated using the independent 30% testing dataset to assess the performance of the final models on observations not used during model training or cross-validation. The same evaluation procedure was applied separately to grain yield and WUE.

2.12. Complete Machine Learning Workflow

The complete machine learning workflow is illustrated in Fig. 6, which summarizes the sequential steps followed in this study. The workflow consisted of four main phases: (1) data collection from field experiments, where growth parameters (plant height, leaf length, and biomass) and yield components (cob length, cob diameter, grains per cob, and water applied) were measured; (2) data preparation, including feature engineering (static and dynamic features), 70/30 data splitting, five-fold cross-validation, and feature scaling; (3) model training using Random Forest, XGBoost, Linear Regression, and Ridge Regression; and (4) model evaluation using R^2 , RMSE, and MAE to compare the predictive performance of the four machine learning models.

Model evaluation using R^2 , RMSE, and MAE was used to compare predictive performance among the four machine learning models. The field experiment results were used separately to evaluate irrigation treatments in terms of grain yield, WUE, and water saving, and to identify a favorable irrigation level for maize production under the semi-arid conditions of the study area. These two evaluations served distinct purposes: machine learning assessed the accuracy of yield and WUE prediction, while the field experiment assessed the agronomic performance of the irrigation treatments.

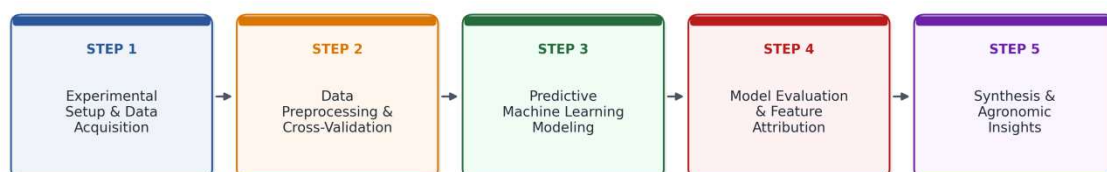


Fig. 6 Methodological framework for predicting maize (*Zea mays* L.) yield and water use efficiency under deficit irrigation.

2.13. Statistical Analysis

All data on maize agronomic parameters, grain yield, and water use efficiency indices were subjected to analysis of variance (ANOVA) using R (version 3.6.3). Treatment means were separated using Fisher's Least Significant Difference (LSD) test at a 5% significance level ($p < 0.05$).

3. Results and discussion

3.1. Effects of Irrigation levels on Growth and Yield

3.1.1. Plant height and Leaf length

Plant height (PH) differed significantly among irrigation treatments ($p < 0.05$), showing a progressive decline with increasing water deficit (Table 4; Fig. 7a). The highest PH was recorded under full irrigation (T1: 100% ET_c) at 289.46 cm, followed by 278.00, 266.00, 245.00, and 231.00 cm under T2 (85%), T3 (70%), T4 (55%), and T5 (40% ET_c), respectively. Compared with full irrigation, severe deficit irrigation (T5) reduced PH by 20.20%, indicating a substantial reduction in maize vegetative growth under severe water stress. These results align with previous studies showing that adequate soil moisture promotes vegetative growth, whereas water deficits limit plant height (Karasu et al., 2015; Yamusa et al., 2017; Kumar et al., 2018; Melkie et al., 2024)

Leaf length (LL) also differed significantly among irrigation treatments ($p < 0.05$), decreasing progressively with increasing water deficit (Table 4; Fig. 7b). Leaf length was highest under full irrigation (T1: 100% ET_c) at 93.00 cm and decreased to 88.00, 86.00, 70.67, and 55.67 cm under T2 (85%), T3 (70%), T4 (55%), and T5 (40% ET_c), respectively. T2 and T3 did not differ significantly, whereas LL declined significantly under the more severe deficit treatments. Compared with T1, T5 reduced LL by 40.14%, demonstrating the strong effect of severe water deficit on leaf development. These results highlight the high sensitivity of maize leaf growth to water deficits, with reductions becoming more pronounced at lower irrigation levels. Similar findings have been reported by Setu et al. (2023), who observed decreased leaf elongation under deficit irrigation. Similarly, Chenu et al. (2008) reported that soil water deficit substantially reduced maize leaf length, demonstrating the sensitivity of leaf expansion to water stress.

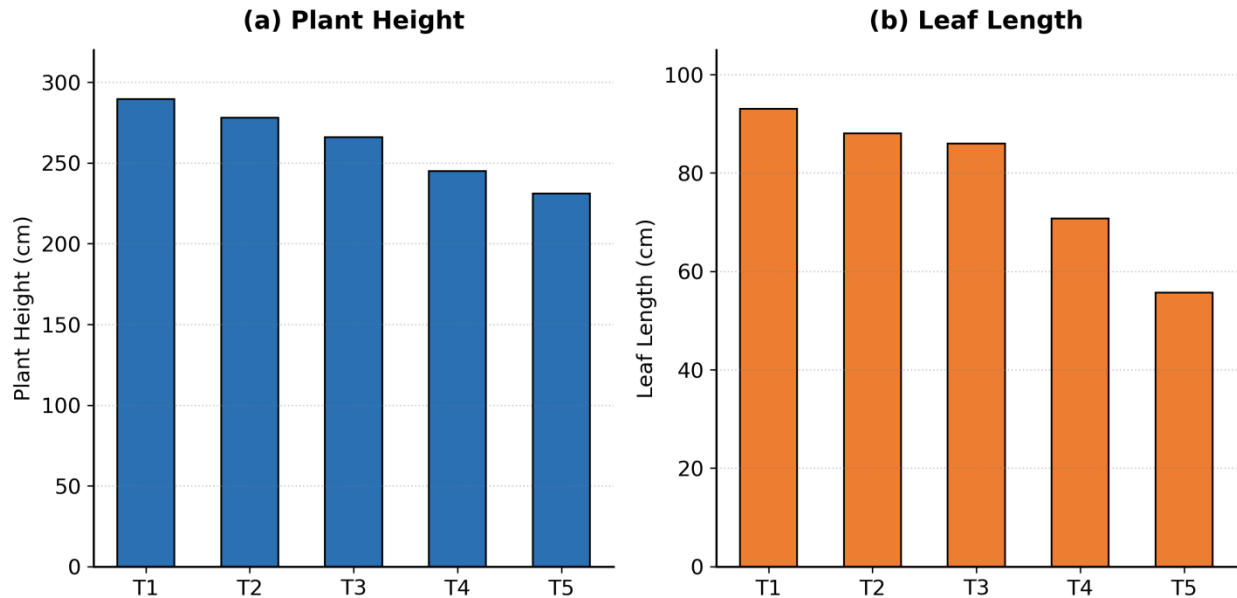


Fig. 7 Response of maize growth parameters (a) plant height (cm) and (b) leaf length (cm)—under deficit irrigation

3.1.2. Above-ground biomass (AGB)

Above-ground biomass (AGB) differed significantly among irrigation treatments ($p < 0.05$), showing a progressive decline with increasing water deficit (Table 4). The highest AGB was recorded under full irrigation (T1: 100% ETc) at 1695.17 g plant⁻¹, followed by 1553.33, 1314.93, 1165.60, and 1036.63 g plant⁻¹ under T2 (85%), T3 (70%), T4 (55%), and T5 (40% ETc), respectively. Compared with full irrigation, severe deficit irrigation (T5) reduced AGB by 38.85%, indicating that increasing water stress substantially limited biomass accumulation.

Similar trends have been reported in previous studies, with increased irrigation enhancing biomass accumulation (Greaves & Wang, 2017; Zhang et al. 2019). Admasu, (2019) indicated that irrigation-induced variability significantly affects biological yield, which is closely associated with above-ground biomass. Agossou Gadédjisso-Tossou et al., (2020) documented maximum biomass under full irrigation, supporting the present findings. Overall, these results underscore the critical role of adequate soil moisture in promoting maize growth and optimizing biomass accumulation.

Table 4 Effects of irrigation levels on maize growth parameters: plant height, leaf length, and above-ground biomass.

Treatment	ETc (%)	PH(cm)	LL(cm)	AGB(g plant ⁻¹)
-----------	---------	--------	--------	-----------------------------

T1	100	289.46 ^a	93 ^a	1695.17 ^a
T2	85	278 ^b	88 ^b	1553.33 ^b
T3	70	266 ^c	86 ^b	1314.93 ^c
T4	55	245 ^d	70.667 ^c	1165.6 ^d
T5	40	231 ^e	55.667 ^d	1036.63 ^e
CV (%)	—	6.51	2.8	6.7
LSD@ (5%)	—	3.2	3.1	17

Note: PH=plant height, LL=leaf length, AGB=above ground biomass. Means within the same column followed by different superscript letters differ significantly at $p < 0.05$ according to the LSD test.

3.2.Effects of Water Levels on Yield, Yield Components, and Harvest Index

3.2.1. Cob Length and Cob Diameter

Cob length and cob diameter showed clear responses to irrigation treatments, with both traits decreasing progressively as irrigation water availability declined (Fig. 8). Full irrigation (100% ETc) produced the longest cobs (20.6 cm), whereas the shortest cobs (13.0 cm) were recorded under severe water deficit (40% ETc) ($p < 0.05$; Table 5; Fig. 8a). Similarly, cob diameter was highest under full irrigation (17.9 cm) and decreased to 10.1 cm under 40% ETc (Fig. 8b). These results demonstrate that reduced water availability adversely affected cob development, resulting in shorter and thinner cobs under water-deficit conditions. The reduction in cob length with decreasing irrigation is consistent with previous findings reported by (Oktem, 2008; Karasu et al., 2015; Aydinsakir et al., (2013). Similarly, (Arash Farsiani, 2011 ; Chinasho et al., 2023) reported reductions in cob length and diameter under water stress, with the highest values generally observed under non-stressed conditions. Overall, the results highlight the importance of adequate irrigation in maintaining cob morphological development and maximizing yield-related traits in maize.

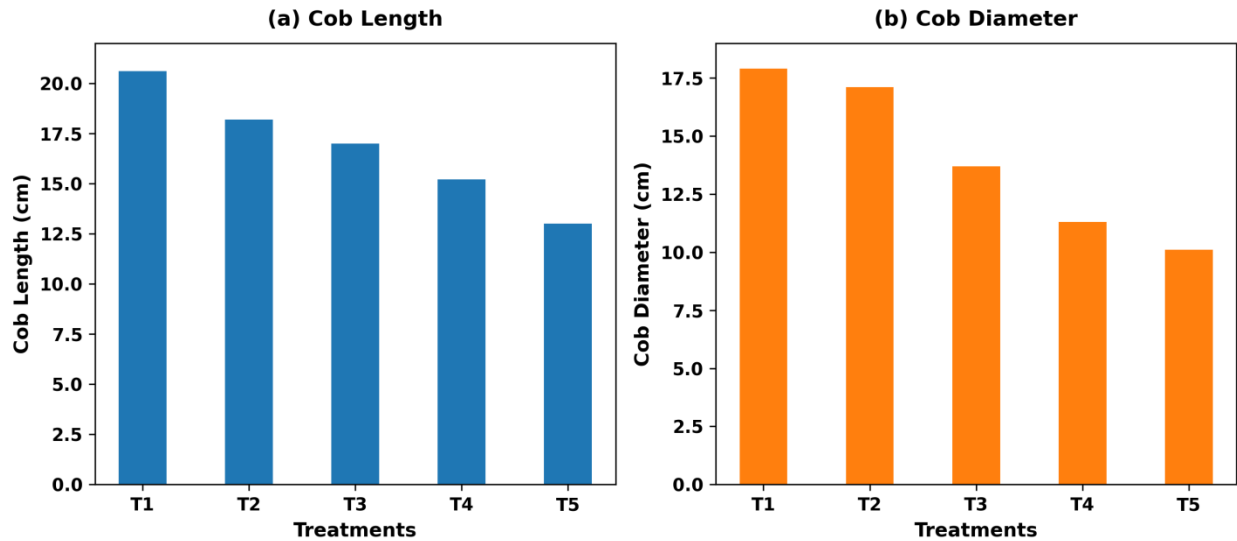


Fig. 8 Effect of irrigation treatments on (a) cob length and (b) cob diameter of maize under different irrigation levels

3.2.2. Number of Grains per Cob

Irrigation level significantly influenced the number of grains per cob ($p < 0.05$; Table 5). The highest mean grain number per cob (667.3) was recorded under full irrigation (100 % ETc) and did not differ significantly from the 85 % ETc treatment. In contrast, the lowest grain number per cob (484.6) occurred under severe water deficit (40 % ETc), showing a clear declining trend as water availability decreased. These results are consistent with (Pandey et al. 2000; Farré & Faci, 2009)), who reported that well-watered maize produces the highest grain counts per cob, whereas water stress causes substantial reductions. Similarly, (Sah & Zamora, 2005) observed the maximum grain number per cob (496.33) under full irrigation, compared with the lowest value (453.66) under rain-fed conditions, further demonstrating the sensitivity of maize reproductive development to water availability.

3.2.3. Harvest Index

Harvest index (HI) was significantly influenced by irrigation level ($p < 0.05$; Table 5). The maximum HI (51%) was observed under the 70% ETc treatment, whereas the lowest value (40%) occurred under severe water stress at 40% ETc. Notably, the response of HI to irrigation level was not strictly linear, reflecting complex interactions between biomass partitioning and water availability. This non-linear response differs from studies reporting maximum harvest index under adequate irrigation, suggesting that assimilate partitioning under deficit irrigation

may vary with irrigation level, crop growth conditions, and management practices (Karasu et al., 2015; (Admasu, 2019; Melkie et al., 2024).

3.2.4. Grain Yield

Maize grain yield was significantly influenced by irrigation treatments ($p < 0.05$; Table 5). Full irrigation (100 % ETc) produced the highest grain yield (7,544 kg.ha⁻¹), whereas the lowest yield occurred under severe deficit irrigation (40 % ETc), declining to 3,556 kg.ha⁻¹. As shown in Fig. 9, Grain yield increased consistently with increasing irrigation level, indicating a strong response of maize productivity to water availability. These findings underscore the importance of optimal irrigation for maximizing yield and suggest that water deficits impose physiological constraints on kernel development and overall crop performance. The observed trends are consistent with previous studies showing that proper soil moisture management through water-balance-based irrigation scheduling enhances maize yield and yield components ((Feres & Soriano, 2006; Igbadun et al., 2008; Genemo et al., 2022; Kumar et al., 2018;; Admasu, 2019).

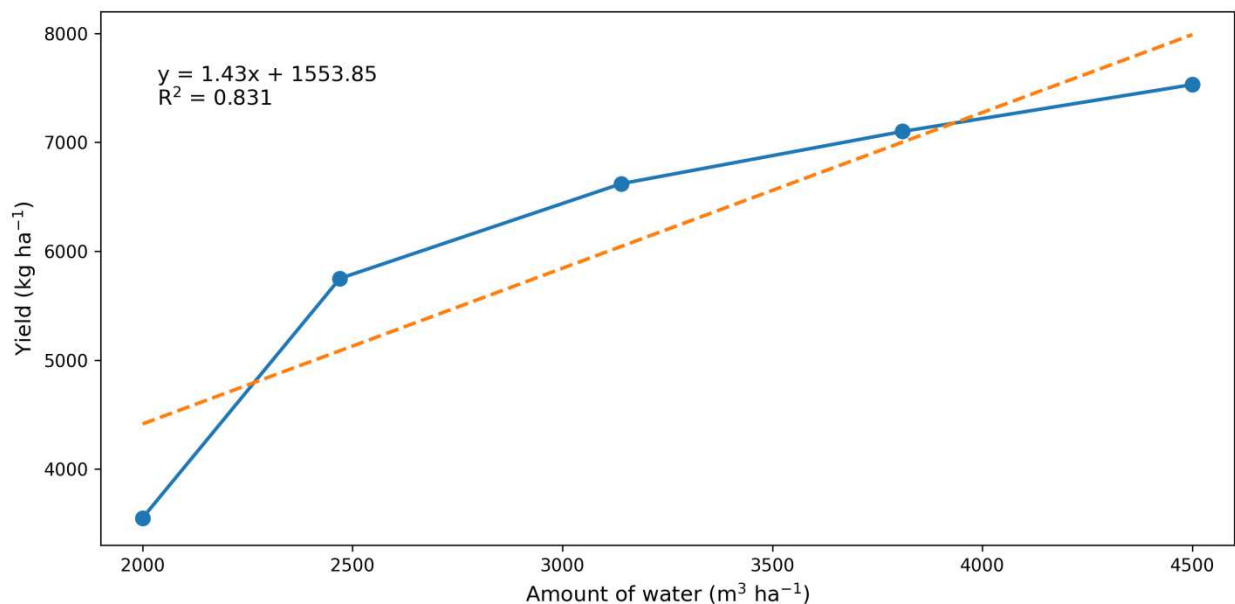


Fig. 9 Relationship between seasonal water use and maize grain yield under varying irrigation regimes

3.2.5. Water Use Efficiency

Water use efficiency (WUE) varied among the irrigation treatments Table (5). WUE increased from 1.68 kg m⁻³ under T1 (100% ETc) to 2.32 kg m⁻³ under T4 (55% ETc), but declined to 1.98 kg m⁻³ under T5 (40% ETc). This indicates that moderate deficit irrigation improved water

productivity, whereas severe water stress reduced WUE. Similar findings were reported for maize in Ethiopia, where deficit irrigation improved water use efficiency compared with full irrigation (Melkie et al., 2024; Gebeyhu Reta et al., 2024).

Although 55% ETc (T4) produced the highest WUE (2.32 kg m⁻³), grain yield declined to 5,746 kg ha⁻¹ compared with 7,544 kg ha⁻¹ under full irrigation. In this study, 70% ETc (T3) produced 6,621 kg ha⁻¹ grain yield and a WUE of 2.10 kg m⁻³, while saving 30% of irrigation water compared with full irrigation. Considering the balance among grain yield, WUE, and irrigation-water saving, 70% ETc (T3) provided a favorable irrigation level under the experimental conditions.

Table 5 Effects of irrigation levels on maize yield components, grain yield, harvest index, and water use efficiencies.

Treatments	ETc (%)	water applied (m ³ ha ⁻¹)	CL(cm)	CD(cm)	NG (cob ⁻¹)	Yield (kg ha ⁻¹)	HI (%)	WUE (kg m ⁻³)
T1	100	4,500	20.6 ^a	17.9 ^a	667.3 ^a	7,544 ^a	48 ^b	1.68
T2	85	3,825	18.2 ^b	17.1 ^a	650.0 ^a	7,100 ^b	46 ^c	1.86
T3	70	3,150	17.0 ^c	13.7 ^b	604.3 ^b	6,621 ^c	51 ^a	2.1
T4	55	2,475	15.2 ^d	11.3 ^c	531.6 ^c	5,746 ^d	43 ^d	2.32
T5	40	1,800	13.0 ^e	10.1 ^c	484.6 ^d	3,556 ^e	40 ^e	1.98
CV(%)	—	—	2	4.4	1.3	0.8	1.5	0.7
LSD(5%)	—	—	0.6	1.1	14.5	94.8	1.3	0.03

Note: CL=cob length, CD=cob diameter, NG=number of cob, HI=harvest index and WUE=water use efficiency. Means within the same column followed by different superscript letters differ significantly at $p < 0.05$ according to the LSD test.

3.3. Amount of Water Saved and Yield Reduction

As presented in Table 6, reducing irrigation from 100% to 40% ETc substantially increased water savings but progressively reduced maize grain yield. Full irrigation (T1) applied 4,500 m³ ha⁻¹ and produced 7,544 kg ha⁻¹. At 85% ETc (T2), water saving was 675 m³ ha⁻¹ (15.0%), with a relatively small yield reduction of 444 kg ha⁻¹ (5.9%). At 70% ETc (T3), water saving increased to 1,350 m³ ha⁻¹ (30.0%), while yield reduction reached 923 kg ha⁻¹ (12.2%). Further reductions to 55% and 40% ETc saved 2,025 m³ ha⁻¹ (45.0%) and 2,700 m³ ha⁻¹ (60.0%), respectively, but caused yield reductions of 1,798 kg ha⁻¹ (23.8%) and 3,988 kg ha⁻¹ (52.9%).

Overall, the results demonstrate a clear trade-off between irrigation-water saving and grain yield. The 70% ETc treatment (T3) provided a favorable balance among grain yield, water-use

efficiency, and irrigation-water saving, producing 6,621 kg ha⁻¹ while saving 1,350 m³ ha⁻¹ (30.0%) of irrigation water compared with full irrigation. In contrast, the more severe deficit treatments resulted in substantially greater yield losses. Thus, under the experimental conditions, 70% ET_c (T3) can be considered a favorable deficit-irrigation level when grain yield, WUE, and water saving are considered together (Admasu, 2019; Kumar et al. 2018).

Table 6 Amount of water saved and yield reduction under deficit irrigation treatments.

Treatm ent	Irrigation Applied (m ³ ha ⁻¹)	Yield (kg ha ⁻¹)	Yield reduction (kg ha ⁻¹)	Yield reduction (%)	Water saved (m ³ ha ⁻¹)	Water Saved (%)
T1	4,500	7,544	—	—	—	—
T2	3,825	7,100	444	5.9	675	15
T3	3,150	6,621	923	12.2	1,350	30
T4	2,475	5,746	1,798	23.8	2,025	45
T5	1,800	3,556	3,988	52.9	2,700	60

3.4. Comparative Performance of Machine Learning Models

Four machine learning models were evaluated using five-fold cross-validation, with final predictive performance assessed on the independent 30% testing dataset Table (7). For grain yield prediction, XGBoost achieved the highest test-set performance ($R^2 = 0.8912$, RMSE = 183.01 kg ha⁻¹, MAE = 155.05 kg ha⁻¹), followed by Random Forest ($R^2 = 0.7489$), Linear Regression ($R^2 = 0.6587$), and Ridge Regression ($R^2 = 0.5231$). The stronger performance of XGBoost indicates that its nonlinear boosting structure captured the relationships between crop-growth characteristics and grain yield more effectively in the present dataset. This finding agrees with (Babaie Sarijaloo et al., 2021) who identified XGBoost as the best-performing model for corn-hybrid yield prediction. However Jeong et al. (2016) reported better performance of Random Forest than multiple linear regressions for maize and other crop-yield predictions, showing that model performance depends on the predictor variables and dataset structure.

For WUE prediction, Ridge Regression produced the highest performance ($R^2 = 0.6540$, RMSE = 0.210 kg m⁻³, MAE = 0.171 kg m⁻³), followed by Linear Regression ($R^2 = 0.5185$), Random Forest ($R^2 = 0.5048$), and XGBoost ($R^2 = 0.3478$). The better performance of Ridge Regression suggests that a regularized linear relationship was more suitable for WUE prediction under the limited-data conditions of this experiment. This finding is consistent with recent work in Southern Ethiopia, where Ridge and Lasso regression performed strongly in predicting the crop yield response factor (Ky) under different irrigation levels Ebstu et al (2026).

Dehghanisanij et al. (2022) likewise demonstrated the applicability of machine learning for WUE and yield prediction, although their study used apple orchards and a different hybrid modeling approach.

Overall, the contrasting performance between the two prediction tasks indicates that model effectiveness depended on the target variable and predictor structure. XGBoost was more effective for grain yield, whereas Ridge Regression provided a better fit for WUE

Table 7 Final predictive performances of machine learning models on the independent testing dataset for maize grain yield and WUE.

Model	Yield R^2	Yield RMSE (kg ha ⁻¹)	Yield MAE (kg ha ⁻¹)	WUE R^2	WUE RMSE (kg m ⁻³)	WUE MAE (kg m ⁻³)
XGB	0.8912	183.01	155.05	0.3478	0.419	0.357
RF	0.7489	251.56	218.88	0.5048	0.252	0.221
LR	0.6587	190.45	172.61	0.5185	0.248	0.234
RR	0.5231	314.65	273.94	0.6540	0.210	0.171

Note: XGB = XGBoost; RF = Random Forest; LR = Linear Regression; RR = Ridge Regression; R^2 = coefficient of determination; RMSE = root mean square error; MAE = mean absolute error.

3.5. Observed and Predicted Grain Yield and Water Use Efficiency

The five observations in the independent test set showed generally close agreement between observed and predicted grain-yield values across the irrigation treatments (Table 8). XGBoost produced predictions close to the observed values, particularly at the 100% and 40% ET_c treatments. At 100% ET_c, XGBoost predicted 7500 kg ha⁻¹ compared with the observed yield of 7544 kg ha⁻¹, corresponding to a deviation of 0.6%. At 40% ET_c, the predicted yield was 3600 kg ha⁻¹ compared with 3556 kg ha⁻¹ observed, representing a deviation of 1.2%. In comparison, Linear Regression overestimated the yield by approximately 10.8% at 40% ET_c. For WUE, the predicted values also showed reasonable agreement with the observed values, although larger deviations were evident under severe deficit irrigation. These results indicate that XGBoost provided close predictions for the observed test-set values under the experimental conditions.

Table 8 Observed and predicted maize grain yield and water use efficiency for the independent test set under different irrigation treatments

Treatments	ET _c (%)	Actual Yield (kg ha ⁻¹)	XGB Yield	RF Yield	LR Yield	RR Yield	Actual WUE (kg m ⁻³)	XGB WUE	RF WUE	LR WUE	RR WUE
T5	40	3,556	3,600	3,650	3,700	3,500	1.98	2.00	2.02	1.97	1.96
T4	55	5,746	5,700	5,650	5,750	5,600	2.32	2.30	2.28	2.31	2.35
T3	70	6,621	6,600	6,550	6,650	6,700	2.10	2.10	2.08	2.10	2.12

Treatments	ETc (%)	Actual Yield (kg ha ⁻¹)	XGB Yield	RF Yield	LR Yield	RR Yield	Actual WUE (kg m ⁻³)	XGB WUE	RF WUE	LR WUE	RR WUE
T2	85	7,100	7,090	6,950	7,150	6,900	1.86	1.86	1.84	1.86	1.88
T1	100	7,544	7,500	7,400	7,550	7,600	1.68	1.68	1.66	1.68	1.70

Note: Predictions are from the final models evaluated on the independent 30% test set ($n = 5$). XGB = XGBoost; RF = Random Forest; LR = Linear Regression; RR = Ridge Regression; WUE = water use efficiency.

Figure 10 illustrates the relationship between observed and predicted grain yield and WUE for the independent test observations. The XGBoost predictions generally followed the observed values, indicating that the model captured the main variation in grain yield across the irrigation treatments. The agreement was particularly evident for the observations under full and severe deficit irrigation. For WUE, the predictions were also generally consistent with the observed pattern, although greater deviations occurred under severe water deficit. However, because the independent test set contained only five observations, the graphical agreement should be interpreted cautiously and does not by itself demonstrate broad predictive generalizability.

The observed–predicted agreement is broadly comparable with previous applications of machine learning to maize yield and WUE prediction. Asamoah et al. (2024) reported a model efficiency coefficient of 0.81 for Random Forest-based maize-yield prediction using 3,136 treatment plots from 482 field trials. Wang et al. (2025) reported test R^2 values of 0.636, 0.536, 0.527, and 0.523 for Random Forest, XGBoost, SVR, and GPR, respectively. Dehghanisani et al. (2022) demonstrated the potential of machine learning for estimating both crop yield and WUE, while Kukal (2024) reported that irrigation depth was an important predictor of maize WUE. However, these studies used substantially larger and more diverse datasets than the present experiment. Therefore, the present results should be considered exploratory, and validation using larger multi-season and multi-location datasets is required.

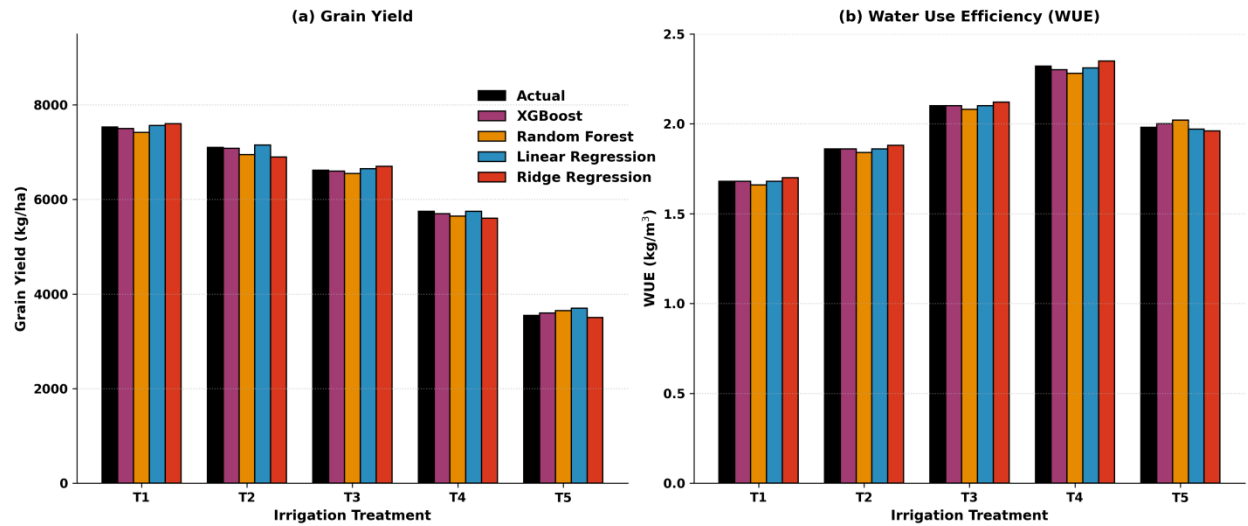


Fig. 10 Comparison of observed and machine learning-predicted maize grain yield and water use efficiency under different irrigation treatments

3.6. Feature Importance Analysis

Feature importance analysis revealed distinct patterns in the relative contributions of plant growth and morphological traits to grain-yield and water-use efficiency (WUE) prediction (Fig. 11a,b). The results demonstrated that the most informative features differed between the two prediction models, highlighting the importance of considering different plant characteristics when predicting crop productivity and water-use performance.

For grain yield (Fig. 11a), biomass growth rate was the most influential feature, accounting for 47.7% of the total importance, followed by plant height (31.6%) and leaf-length AUC (16.5%). Collectively, these three features represented 95.8% of the total importance, demonstrating their dominant role in grain-yield prediction. In contrast, plant height growth rate and cob length contributed only 2.2% and 2.0%, respectively. The predominance of biomass growth rate highlights the importance of crop growth dynamics in predicting final grain production. This finding is consistent with Babaie Sarijaloo et al. (2021) who demonstrated the relevance of crop-growth characteristics in machine-learning-based crop-yield prediction.

In comparison, WUE prediction was primarily driven by cob-related morphology and accumulated growth characteristics (Fig. 11b). Cob shape ratio emerged as the most influential feature, contributing 43.6% of the total importance, followed by plant height (20.2%) and biomass AUC (17.2%). Together, these three features accounted for 81.0% of the total importance. Cob length contributed an additional 9.7%, whereas biomass growth rate (2.7%), PH

AUC (2.3%), PH growth rate (1.8%), LL AUC (1.5%), and LL growth rate (1.1%) had relatively minor contributions. The prominence of cob shape ratio and cob length suggests that reproductive morphological characteristics provided substantial predictive information for WUE under the evaluated conditions. This observation is in line with Elbeltagi et al. (2022), who highlighted the potential of machine-learning approaches to predict crop water productivity under different water-management conditions.

Overall, the feature importance profiles revealed a clear distinction between the two prediction models. Grain-yield prediction was predominantly associated with biomass accumulation dynamics and plant stature, whereas WUE prediction relied more strongly on cob morphology, plant height, and accumulated biomass. Notably, plant height was an important common feature in both grain yield and WUE, contributing 31.6% to grain-yield prediction and 20.2% to WUE prediction. These contrasting importance patterns indicate that the traits providing the greatest predictive information for grain production were not necessarily those providing the greatest information for WUE, reflecting the different growth and morphological attributes captured by the two models.

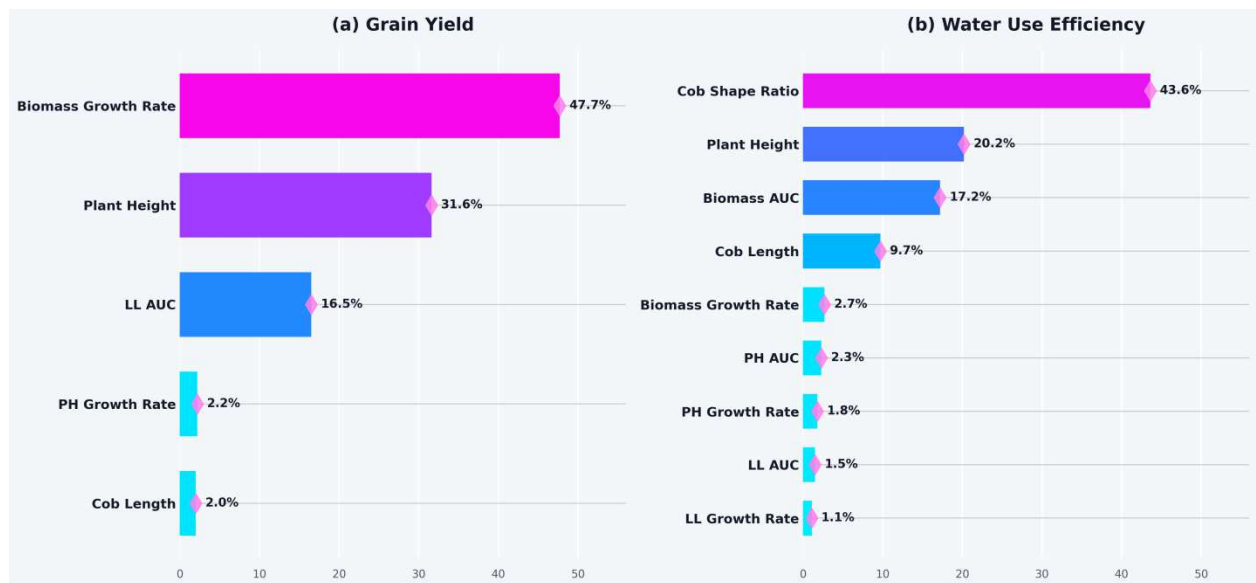


Fig. 11 Relative importance (%) of plant growth and morphological traits for predicting (a) grain yield and (b) water use efficiency (WUE).

4. Conclusion

This study showed the potential of machine learning to predict maize grain yield and water use efficiency (WUE) under water-balance-based deficit irrigation in semi-arid Ethiopia. XGBoost showed the highest test-set performance for grain-yield prediction, whereas Ridge Regression showed the highest test-set performance for WUE prediction within the evaluated dataset. Temporal crop-growth features, particularly biomass growth rate and plant height, were important predictors of grain yield, while cob shape ratio was the most important predictor of WUE. Field results showed that deficit irrigation increased irrigation-water savings but generally reduced grain yield. Among the evaluated treatments, 70% ET_c (T3) provided a favorable balance among grain yield, WUE, and irrigation-water savings, producing 6,621 kg ha⁻¹ grain yield and 2.10 kg m⁻³ WUE while saving 30% of irrigation water compared with full irrigation. However, the small number of experimental observations and the five-observation independent test set limit the generalizability of the machine-learning results. Therefore, larger multi-season and multi-location datasets, together with independent validation, are needed to assess the robustness and broader applicability of the models.

Declarations

Funding

The authors received no specific funding for this research.

Competing interests

The authors declare that they have no competing interests.

Ethics approval

Not applicable. This study involved field experimentation on maize and did not involve human participants or animals.

Consent to Participate

Not applicable. This study did not involve human participants.

Consent for publication

Not applicable.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Author Contributions

Melaku Adugnaw Walle designed the experiment, conducted the field measurements, performed the data analysis, developed and evaluated the machine learning models, and drafted the manuscript. Hailemariam Molla Ashagre contributed to the study design, supervised the research, reviewed and critically revised the manuscript, and approved the final version.

References

- Admasu R, Michael AW, Hordofa T (2019) Effect of moisture stress on maize (*Zea mays* L.) yield and water productivity. *Int J Environ Sci Nat Res* 16(4):83–87. <https://doi.org/10.19080/IJESNR.2019.16.555945>
- Abedinpour, M. (2015). Evaluation of growth-stage-specific crop coefficients of maize using weighing lysimeter. *Soil and Water Research*, 10(2), 99–104. <https://doi.org/10.17221/63/2014-SWR>
- Agyeman, B. T., et al. (2025). *A semi-centralized multi-agent reinforcement learning framework for efficient irrigation scheduling*. *Control Engineering Practice*, 155, 106183. <https://doi.org/10.1016/j.conengprac.2024.106183>
- Asamoah E, Heuvelink GBM, Chairi I, Bindraban PS, Logah V (2024) Random forest machine learning for maize yield and agronomic efficiency prediction in Ghana. *Heliyon* 10(17):e37065. <https://doi.org/10.1016/j.heliyon.2024.e37065>
- Assefa, B. T., Chamberlin, J., Reidsma, P., Silva, J. V., & van Ittersum, M. K. (2020). Unravelling the variability and causes of smallholder maize yield gaps in Ethiopia. *Food Security*, 12, 83–103. <https://doi.org/10.1007/s12571-019-00981-4>
- Aydinsakir K, Erdal S, Buyuktas D, Bastug R, Toker R (2013) The influence of regular deficit irrigation applications on water use, yield, and quality components of two corn (*Zea mays* L.) genotypes. *Agric Water Manag* 128:65–71. <https://doi.org/10.1016/j.agwat.2013.06.013>
- Babaie Sarijaloo F, Porta M, Taslimi B, Pardalos PM (2021) Yield performance estimation of corn hybrids using machine learning algorithms. *Artif Intell Agric* 5:82–89. <https://doi.org/10.1016/j.aiia.2021.05.001>
- Ben Hassen T, El Bilali H, Allahyari MS, Baya Chatti C (2025) Traditional irrigation knowledge for sustainable water resource management in arid environments: Insights from the MENA region. *J Arid Environ* 231:105466. <https://doi.org/10.1016/j.jaridenv.2025.105466>
- Birhanu B, Kebede S, Charles K, Taye M, Atlaw A, Birhane M (2021) Impact of natural and anthropogenic stresses on surface and groundwater supply sources of the Upper Awash sub-basin, central Ethiopia. *Front Earth Sci* 9:656726. <https://doi.org/10.3389/feart.2021.656726>
- Chenu, K., Franck, N., Dauzat, J., Barczi, J.-F., Rey, H., & Lecoeur, J. (2008). Short-term responses of leaf growth rate to water deficit scale up to whole-plant and crop levels: An integrated modelling approach in maize. *Plant, Cell & Environment*, 31(3), 378–391. <https://doi.org/10.1111/j.1365-3040.2007.01772.x>
- Chinasho A, Bedadi B, Tefera TL, Tana T, Hordofa T, Elias B (2023) Response of maize to irrigation and blended fertilizer levels for climate smart food production in Wolaita Zone, southern Ethiopia. *J Agric Food Res* 12:100551. <https://doi.org/10.1016/j.jafr.2023.100551>
- Chlingaryan A, Sukkarieh S, Whelan B (2018) Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput Electron Agric* 151:61–69. <https://doi.org/10.1016/j.compag.2018.05.012>

- Colusse GA, Borges-Mendes CR, Rabello-Duarte ME, de Carvalho JC, Noseda MD (2020) Effects of different culture media on physiological features and laboratory scale production cost of *Dunaliella salina*. *Biotechnol Rep* 27:e00508. <https://doi.org/10.1016/j.btre.2020.e00508>
- Dehghanisani H, Emami H, Emami S, Rezaverdinejad V (2022) A hybrid machine learning approach for estimating the water-use efficiency and yield in agriculture. *Sci Rep* 12:6728. <https://doi.org/10.1038/s41598-022-10844-2>
- Ebstu E. T., Hatiye S. D., & Goshime, D. W. (2026). Integrating machine learning and field experiments to model crop coefficient and yield response of lowland irrigated wheat in Southern Ethiopia. *International Journal of Agronomy*, 2026, Article 7370433. <https://doi.org/10.1155/iaa/7370433>
- Elbeltagi, A., Srivastava, A., Kushwaha, N. L., Juhász, C., Tamás, J., & Nagy, A. (2023). Meteorological data fusion approach for modeling crop water productivity based on ensemble machine learning. *Water*, 15(1), 30. <https://doi.org/10.3390/w15010030>
- El-Rawy M, Batelaan O, Al-Arifi N, Alotaibi A, Abdalla F, Gabr ME (2023) Climate change impacts on water resources in arid and semi-arid regions: A case study in Saudi Arabia. *Water* 15(3):606. <https://doi.org/10.3390/w15030606>
- Farré I, Faci JM (2009) Deficit irrigation in maize for reducing agricultural water use in a Mediterranean environment. *Agric Water Manag* 96(3):383–394. <https://doi.org/10.1016/j.agwat.2008.07.002>
- Farsiani A, Ghobadi M-E, Jalali-Honarmand S (2011) The effect of water deficit and sowing date on yield components and seed sugar contents of sweet corn (*Zea mays* L.). *Afr J Agric Res* 6(26):5769–5774. <https://doi.org/10.5897/AJAR11.1242>
- Fereres E, Soriano MA (2007) Deficit irrigation for reducing agricultural water use. *J Exp Bot* 58(2):147–159. <https://doi.org/10.1093/jxb/erl165>
- Gadédjisso-Tossou A, Avellán T, Schütze N (2020) Impact of irrigation strategies on maize (*Zea mays* L.) production in the savannah region of northern Togo (West Africa). *Water SA* 46(1):141–152. <https://doi.org/10.17159/wsa/2020.v46.i1.7894>
- Gebeyhu Reta, B., Dagalo Hatiye, S., & Muluneh Finsa, M. (2024). Assessment of irrigation water management performance indicators and mitigation measure in Arba Minch irrigation scheme, Ethiopia. *Advances in Agriculture*, 2024, Article 5596514. <https://doi.org/10.1155/2024/5596514>
- Gebreselassie, Y., Ayana, M., & Tadele, K. (2015). Field experimentation based simulation of yield response of maize crop to deficit irrigation using AquaCrop model, Arba Minch, Ethiopia. *African Journal of Agricultural Research*, 10(4), 269–280. <https://doi.org/10.5897/AJAR2014.8703>
- Genemo G, Mokonen E, Bedane H (2022) Effect of deficit irrigation on yield and water use efficiency of maize at Indris irrigation scheme, Western Oromia, Ethiopia. *Int J Agric Nutr* 4(2):119–123. <https://doi.org/10.33545/26646064.2022.v4.i2b.103>
- Greaves GE, Wang Y-M (2017a) Effect of regulated deficit irrigation scheduling on water use of corn in southern Taiwan tropical environment. *Agric Water Manag* 188:115–125. <https://doi.org/10.1016/j.agwat.2017.04.008>
- Howell TA (2001) Enhancing water use efficiency in irrigated agriculture. *Agron J* 93(2):281–289. <https://doi.org/10.2134/agronj2001.932281x>
- Igbadun HE, Salim BA, Tarimo AKPR, Mahoo HF (2008) Effects of deficit irrigation scheduling on yields and soil water balance of irrigated maize. *Irrig Sci* 27(1):11–23. <https://doi.org/10.1007/s00271-008-0117-0>
- Intergovernmental Panel on Climate Change (2023) Climate change 2022: Impacts, adaptation and vulnerability. Cambridge University Press. <https://doi.org/10.1017/9781009325844>

- Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K.-M., Gerber, J. S., Reddy, V. R., & Kim, S.-H. (2016). Random forests for global and regional crop yield predictions. *PLOS ONE*, 11(6), e0156571. <https://doi.org/10.1371/journal.pone.0156571>
- Kukul MS (2024) Consumptive water use and efficiency of irrigated U.S. corn: Learning from hundreds of site-management-year observations. *Water Resour Res* 60(8):e2024WR037434. <https://doi.org/10.1029/2024WR037434>
- Kumar V, Chandra R, Jain SK (2018) Effect of deficit irrigation in Rabi maize for crop growth, yield, biomass and water use efficiency in North Bihar condition. *Int J Curr Microbiol Appl Sci* 7(6):2694–2701. <https://doi.org/10.20546/ijcmas.2018.706.318>
- Lee SO, Jung Y (2018) Efficiency of water use and its implications for a water-food nexus in the Aral Sea Basin. *Agric Water Manag* 207:80–90. <https://doi.org/10.1016/j.agwat.2018.05.014>
- Lobell DB, Bänziger M, Magorokosho C, Vivek B (2011) Nonlinear heat effects on African maize as evidenced by historical yield trials. *Nat Clim Change* 1:42–45. <https://doi.org/10.1038/nclimate1043>
- McDonald RI, Green P, Balk D, Fekete BM, Revenga C, Todd M, Montgomery M (2011) Urban growth, climate change, and freshwater availability. *Proc Natl Acad Sci USA* 108(15):6312–6317. <https://doi.org/10.1073/pnas.1011615108>
- Melkie T, Jemberu W, Bitew A (2024) Optimizing water use efficiency in maize (*Zea mays* L.) production through deficit irrigation in Gazhen-Fuafuat Kebele, Northwest Ethiopia. *Front Agron* 6:1490423. <https://doi.org/10.3389/fagro.2024.1490423>
- Mondschein Z, Paliwal A, Sida TS, Chamberlin J, Wang R, Jain M (2024) Mapping field-level maize yields in Ethiopian smallholder systems using Sentinel-2 imagery. *Remote Sens* 16(18):3451. <https://doi.org/10.3390/rs16183451>
- Morante-Carballo F, Montalván-Burbano N, Quiñonez-Barzola X, Jaya-Montalvo M, Carrión-Mero P (2022) What do we know about water scarcity in semi-arid zones? A global analysis and research trends. *Water* 14(17):2685. <https://doi.org/10.3390/w14172685>
- Muharomah R, Setiawan BI, Sands GR, Juliana IC, Gunawan TA (2025) A review on enhancing water productivities adaptive to the climate change. *J Water Clim Change* 16(3):860–887. <https://doi.org/10.2166/wcc.2025.240>
- Ntouros K, Papatheodorou K, Gkologkinas G, Drimzakas-Papadopoulos V (2025) A Python framework for crop yield estimation using Sentinel-2 satellite data. *Earth* 6(1):15. <https://doi.org/10.3390/earth6010015>
- Oktem A (2008) Effect of water shortage on yield, and protein and mineral compositions of drip-irrigated sweet corn in sustainable agricultural systems. *Agric Water Manag* 95(9):1003–1010. <https://doi.org/10.1016/j.agwat.2008.03.006>
- Pandey RK, Maranville JW, Admou A (2000) Deficit irrigation and nitrogen effects on maize in a Sahelian environment: I. Grain yield and yield components. *Agric Water Manag* 46(1):1–13. [https://doi.org/10.1016/S0378-3774\(00\)00073-1](https://doi.org/10.1016/S0378-3774(00)00073-1)
- Pereira LS, Allen RG, Smith M, Raes D (2015) Crop evapotranspiration estimation with FAO56: Past and future. *Agric Water Manag* 147:4–20. <https://doi.org/10.1016/j.agwat.2014.07.031>
- Pragg B, Deepagoda TTKC, Cameron K, Di H, Clough TJ, Carrick S (2024) Irrigation scheduling needs to consider both plant-available water and soil aeration requirements. *Vadose Zone J* 23:e20344. <https://doi.org/10.1002/vzj2.20344>

- Sah S. K., & Zamora, O. B. (2005). Effect of water deficit at vegetative and reproductive stages of hybrid, open pollinated variety and local maize (*Zea mays* L.). *Journal of the Institute of Agriculture and Animal Science*, 26, 37–42. <https://doi.org/10.3126/jiaas.v26i0.609>
- Setu T, Legese T, Teklie G, Gebeyhu B (2023) Effect of furrow irrigation systems and irrigation levels on maize agronomy and water use efficiency in Arba Minch, Southern, Ethiopia. *Heliyon* 9(7):e17833. <https://doi.org/10.1016/j.heliyon.2023.e17833>
- Shahhosseini M, Hu G, Huber I, Archontoulis SV (2021) Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. *Sci Rep* 11:1606. <https://doi.org/10.1038/s41598-020-80820-1>
- Tace Y, Tabaa M, Elfilali S, Leghris C, Bensag H, Renault E (2022) Smart irrigation system based on IoT and machine learning. *Energy Rep* 8:1025–1036. <https://doi.org/10.1016/j.egyr.2022.07.088>
- Van Klompenburg T, Kassahun A, Catal C (2020) Crop yield prediction using machine learning: A systematic literature review. *Comput Electron Agric* 177:105709. <https://doi.org/10.1016/j.compag.2020.105709>
- Wang, S., Wang, G., Wang, Y., Wang, Y., Wang, H., Li, W., Wei, J., & Jiang, X. (2025). Integrating meteorological and breeding data to predict maize yields using machine learning algorithms. *Frontiers in Plant Science*, 16, 1722068. <https://doi.org/10.3389/fpls.2025.1722068>
- Yamusa AM, Okogbue EC, Agele SO (2017) Effect of deficit irrigation levels on yield and water productivity of some maize varieties in a semi-arid environment. *Int J Plant Soil Sci* 20(4):1–9. <https://doi.org/10.9734/IJPSS/2017/32847>
- Yigezu TT, Narayanan K, Hordof T (2016) Effect of furrow length and flow rate on irrigation performances and yield of maize. *Int J Eng Res Technol* 5(4). <https://doi.org/10.17577/IJERTV5IS040846>
- Zhang G, Shen D, Ming B, Xie R, Jin X, Liu C, Hou P, Xue J, Chen J, Zhang W, Liu W, Wang K, Li S (2019) Using irrigation intervals to optimize water-use efficiency and maize yield in Xinjiang, northwest China. *Crop J* 7(3):322–334. <https://doi.org/10.1016/j.cj.2018.10.008>
- Zhao L, Zhao X, Pan X, Shi Y, Qiu Z, Li X, Xing X, Bai J (2022) Prediction of daily reference crop evapotranspiration in different Chinese climate zones: Combined application of key meteorological factors and Elman algorithm. *J Hydrol* 610:127822. <https://doi.org/10.1016/j.jhydrol.2022.127822>
- Zou Y, Saddique Q, Ali A, Xu J, Khan MI, Qing M, Azmat M, Cai H, Siddique KHM (2021) Deficit irrigation improves maize yield and water use efficiency in a semi-arid environment. *Agric Water Manag* 243:106483. <https://doi.org/10.1016/j.agwat.2020.106483>