Smart Agricultural Technology 13 (2026) 101870

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Predicting smallholder maize yield using sentinel-2-derived phenological metrics

Wonga Masiza [a,*], Basani Lammy Nkuna [a], Phathutshedzo Eugene Ratshiedana [a], Akhona Madasa [a,b], Lwandile Nduku [a], Tumelo Shwatja [a,b], Johannes George Chirima [a,b], Adolph Nyamugama [a], Khaled Abutaleb [a], Pitso Walter Khoboko [b,c], Hamisai Hamandawana [d]

[a] Agricultural Research Council-Natural Resource and Engineering, Geoinformatics division, Arcadia 0083, South Africa
[b] Centre for Geoinformation Science, Department of Geography, Geoinformatics, and Meteorology, University of Pretoria, Pretoria 0001, South Africa
[c] Neria Solutions, 21354, Turflaagte 2, Bloemfontein, 9323, South Africa
[d] 2990 Mageza St, Protea North, Soweto, South Africa

ARTICLE INFO

Keywords:
Smallholder
Maize
Crop yield
Remote sensing
Phenology

ABSTRACT

Accurate estimation of smallholder crop yields remains challenging due to limited field data and diverse agronomic practices. This study evaluated Sentinel-2-derived phenological metrics for maize yield estimation across three seasons (2018–2019, 2021–2022, and 2024–2025) in a smallholder farming area. Time series vegetation indices processed in Google Earth Engine, smoothed using the Whittaker filter, and transformed into phenological metrics were compared with crop-cut yield data collected from 123 farms. Regularised linear and ensemble models were assessed using cross-validation, with leave-one-season-out validation used to evaluate temporal generalization. Ridge regression achieved higher seasonal accuracy ($R^2 = 0.56$; RMSE = 1148 kg ha^{-1}) confirming that a compact, regularised linear model can be effective with limited data. When data from all seasons were merged, Elastic Net achieved the highest accuracy among the tested models ($R^2 = 0.52$; RMSE = 1401 kg ha^{-1}). SHAP and permutation analyses identified consistent predictors across years, notably pre-peak and cumulative integrals of CIrededge and MCARI, highlighting the importance of early-season canopy development and season-long spectral signals associated with chlorophyll dynamics. Parsimonious Ridge and GB models trained with Top-10 features retained comparable accuracy ($R^2 = 0.49$–0.56). Residual analysis showed negligible bias (mean residuals = -38 to -42 kg ha^{-1}) and moderate dispersion (SD = 1348 to 1448 kg ha^{-1}), although prediction uncertainty remained substantial across the yield range. The findings demonstrate that compact, interpretable models based on satellite-derived phenological metrics can potentially contribute to scalable, data-driven assessments of yield variability in smallholder systems.

1. Introduction

Agriculture is one of the key sectors expected to contribute towards the realization of some of the Sustainable Development Goals (SDGs) espoused by the United Nations [1]. Out of the 17 UN SDGs, SDG 1 and SDG 2 are particularly important for the eradication of poverty and hunger through agriculture. Within this framework, Target 2.3 of SDG 2 explicitly calls for doubling agricultural productivity and incomes of small-scale food producers by 2030 [2]. Achieving these targets requires strategies that can increase crop yields without expanding cropland at the expense of natural ecosystems [3–5]. Alongside crop yield improvements, there is also a need for robust crop monitoring and yield tracking tools. Such tools are particularly important for smallholder farming communities, where productivity is not only poorly documented [6–9] but also far below that of large-scale commercial farms [1].

Since crop yield is a key metric for evaluating agricultural performance, crop yield estimation systems can be used to track smallholder productivity and assess the effectiveness of farmer support programmes. Defined as the quantity of crop harvested per unit area, crop yield can be estimated before or after harvest depending on the purpose of the assessment and available resources [10]. This can be done using a range of methods including traditional crop cut methods [11,10,12], process-based crop models [13,14], statistical and machine learning

* Corresponding author.
E-mail address: masizaw@arc.agric.za (W. Masiza).

https://doi.org/10.1016/j.atech.2026.101870
Received 4 November 2025; Received in revised form 7 February 2026; Accepted 11 February 2026
Available online 12 February 2026

approaches [15,16], as well as methods that use remotely sensed data [17,18]. Each approach has distinct advantages and constraints, with suitability determined by the spatial scale of analysis, the availability of input data, and the degree of field-level detail required for reliable yield estimation.

Although traditional crop cut methods and process-based models are valuable, they face practical constraints in many smallholder farming contexts. Crop cuts can be labour-intensive, costly, and limited in spatial coverage [19], whereas process-based models require detailed input data (e.g., soil properties, weather, and management practices) that are often unavailable at the required temporal and spatial scales [20,14]. It is for these reasons that empirical models and remote sensing-based approaches have been increasingly gaining traction [9,21]. Empirical models leverage readily available historical datasets to establish statistical relationships between observed variables and crop yield outcomes [16,22,23], while remote sensing enables consistent, large-area monitoring of crop growth conditions [17,24]. These approaches provide a complementary framework capable of achieving accurate and operationally feasible yield estimations.

In remote sensing-based crop monitoring, phenology has emerged as a critical factor for understanding crop growth dynamics and improving yield estimation [25–28]. Crop phenology refers to the biological growth and development stages of crops, commonly described using standardized frameworks such as the BBCH scale and observed though in-situ measurements [29,30]. The development stages of crops span key phases of crop establishment, reproduction, and biomass accumulation that are known to exert strong controls on yield formation, and are therefore closely linked to crop vigour, stress response, and final yield [31,32,28]. In contrast, land surface phenology (LSP) represents the spatio-temporal development of vegetation as observed by satellite sensors, capturing integrated canopy-level responses rather than discrete physiological growth stages [33,34]. From satellite-derived vegetation index time series, phenological metrics such as the timing of green up [35], maximum greenness [36], and senescence [37], are extracted and used as proxies for seasonal crop dynamics. Although LSP-derived metrics do not necessarily correspond directly to specific crop development stages, they provide a consistent and scalable means of characterizing crop growth patterns, making them particularly valuable in heterogenous smallholder systems where field-based phenological observations are scarce.

However, the use of satellite-derived phenological information is not without challenges. Sensors such as MODIS, AVHRR, and VIIRS offer long-term and frequent coverage but have coarse spatial resolutions with limited applications in fragmented smallholder landscapes [38–41, 36]. Other systems with finer resolutions often face trade-offs in temporal coverage, temporal resolution, data continuity, data cost, or cloud contamination [42–44]. The Sentinel-2 satellite constellation offers a potential means of addressing the limitations faced by the above-mentioned satellite systems by providing freely available multispectral imagery at relatively high spatial resolution (10–20 m) and short revisit interval [45], making it particularly suitable for monitoring heterogeneous and fragmented agricultural landscapes typical of smallholder farming systems.

On the other hand, while satellite-derived phenological metrics have been increasingly used for crop yield estimation, existing studies suggest that not all such metrics are equally effective, with some failing to capture critical stages of crop development or exhibiting sensitivity to noise, cloud contamination, and temporal data gaps [46,47]. Consequently, many yield prediction studies rely on a limited subset of phenological descriptors derived from one or a small number of vegetation indices, often selected a priori rather than through systematic evaluation. Recent work has explored phenological metrics derived from individual indices or a small set of vegetation indices [26–28,48,49], but comprehensive assessments of a broader, multi-index phenological feature space remain limited, particularly at the field scale in heterogenous smallholder systems. Moreover, relatively few studies have explicitly examined the temporal stability and cross-season transferability of phenological metrics using cross-season evaluation approaches.

Against this background, this study systematically evaluates a wide range of phenological metrics derived from multiple Sentinel-2 vegetation indices for smallholder maize yield estimation in South Africa's OR Tambo District Municipality (ORTDM). Specifically, the study (i) assesses per-season yield prediction performance to characterise intra-season variability, (ii) evaluates cross-season generalization and temporal transferability using leave-one-season-out validation, and (iii) identifies stable and informative phenological predictors across seasons to support parsimonious modelling. By integrating multi-index phenological metrics, feature-stability analysis, and cross-season validation, this work extends existing yield-prediction studies towards more interpretable frameworks suited to data-limited smallholder systems.

2. Materials and methods

2.1. Study area

The study was conducted in the ORTDM of the Eastern Cape Province, South Africa (Fig. 1). The sampled fields were sparsely distributed across three local municipalities: King Sabata Dalindyebo, Mhlontlo, and Nyandeni. Combined, these three local municipalities cover an area of approximately 8327 km^2 of the former Transkei homeland.

Although the area is characterised by variable climate conditions [50,51], recurrent pests and diseases [52–55] and agricultural decline [56–58], agriculture remains the main economic sector with a significant growth potential [59]. Maize is the most cultivated crop with highly variable management practices and farm sizes, which span from less than one hectare to over 100 hectares [60]. Based on the yield surveys conducted in this study and other surveys [61], NPK fertilizer application rates typically range from approximately 50 to 250 kg ha^{-1}, reflecting differences in input access and farmer resources. Multiple maize cultivars are grown across farms and planting dates vary widely from early October to late January depending on rainfall onset and mechanization availability, with some farmers using owned equipment while others rely on hired services [61]. These factors, combined with inter-annual variability in rainfall and temperature, contribute to pronounced differences in growing conditions across fields and seasons. Yellow maize is more dominant and primarily used as livestock feed, while white maize serves as the main staple food for humans across the country [61]. Other crops grown in the area include soybean, sugar bean, potatoes, cabbages, and spinach [54].

Most agricultural land is either owned by the government or managed under the custodianship of tribal authorities and administered as communal land [59]. The district spans a rich mosaic of vegetation types, shaped by its coastal location, varied topography, and climatic gradient. The inland areas are mostly covered by grasslands and savanna with some sections along the coast featuring dense thicket and woody species [62]. The climate is subtropical and humid along the coastal areas, and temperate to moderately dry in the inland [50]. The Agricultural Research Council (ARC)'s agroclimate databank [63] shows that most precipitation occurs between October and March with annual rainfall of 800–1200 mm and 600–900 mm along the coastal belt and the inland plateau, respectively. The coastal regions have mild to warm temperatures year-round with average highs of 24–30 °C in summer and average lows of 8–12 °C in winter. The inland parts have more temperature variation with summer highs of up to 30 °C and winter lows of 3–5 °C.

2.2. Data

The data used in this study consisted of in-situ yield measurements and vegetation indices derived from the Sentinel-2 constellation of satellites, which are part of Europe's Copernicus programme [45].

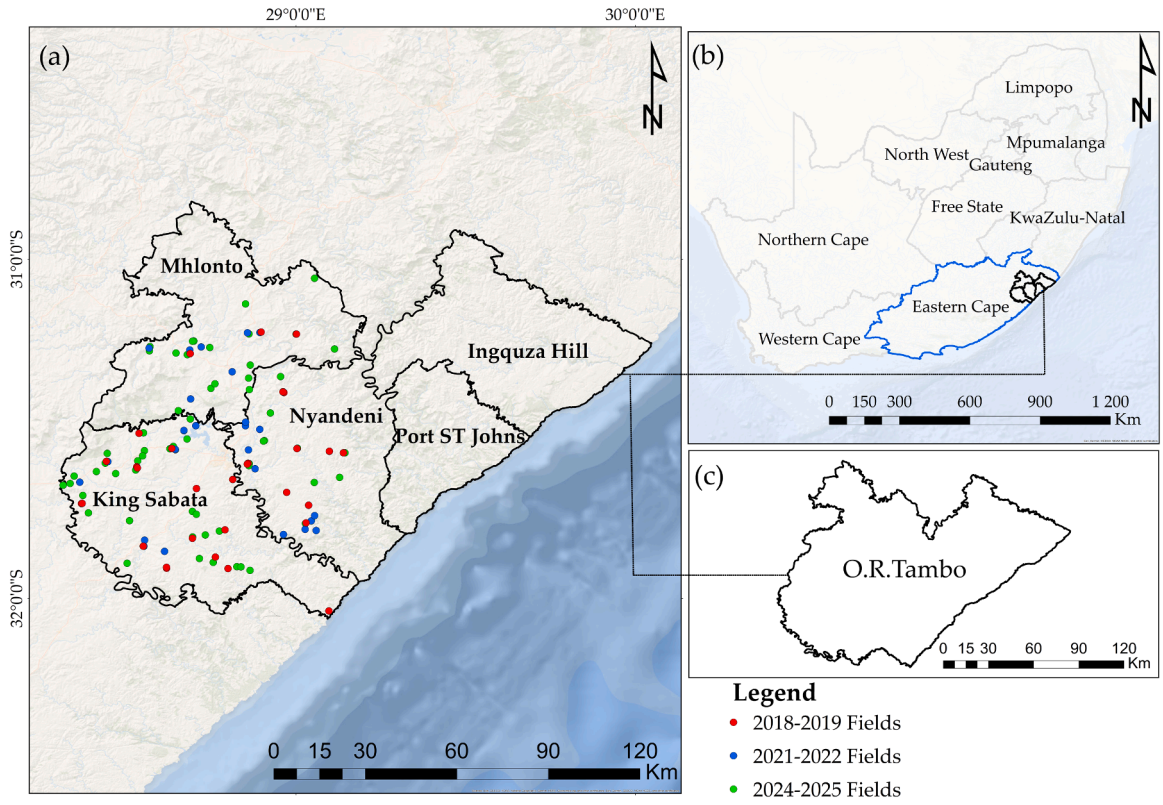

Fig. 1. The OR Tambo District Municipality in Eastern Cape, South Africa.

2.2.1. Yield data

Yield data were collected from 123 farms over three seasons comprising 2018–2019 and 2021–2022 during which 30 farms were included for each of these seasons, and the 2024–2025 season during which 63 farms were covered. Yield surveys were conducted in May and June coinciding with the harvest season, from fields where harvesting had not yet started. These surveys were carried out using the objective yield survey (OYS) method for maize, which is the standard method used by South Africa's Crop Estimates Consortium (CEC) for large-scale commercial farms due to its established abilities to provide reliable data on crop yields through direct measurements of plant characteristics [64].

This method involves first measuring the average spacing of rows (i.e., row-to-row spacing) in a field after which, five representative sampling points are selected following instructions provided in a random steps table. This table provides instructions on how to randomly select points by counting the number of steps to walk along and across the maize rows. At each sampling point, which is a 10-meter section of the sampled row, the surveyor records the number of plants and ears and then harvests a fixed number of ears. The harvested ears are arranged by size, from largest to smallest, and the median ear selected as the representative sample. This process is repeated across all five sampling points in the field.

The selected ears are shelled, and the grain analysed for moisture content and grain mass using a grain moisture meter for the former and a weighing balance scale for the latter. These measurements, along with farm and field details were recorded using standardized survey forms and a mobile application developed by the ARC in partnership with Neria Solutions, which enabled GPS-linked data capture and automated yield calculation. The mathematical formulation for yield calculation is shown in Eqs. (1) and (2):

$$Y_i = 0.95 \times \frac{E_i \times M_i \times \frac{100-H_i}{100-12.5}}{L \times R} \times \frac{10000}{1000} \quad (1)$$

where:

- Y_i = yield at sampling point i (kg ha^{-1})
- E_i = number of ears at sampling point i
- M_i = grain mass at sampling point i (in grams)
- H_i = grain moisture (%) at sampling point i
- L = sampling row length (in meters), which is 10 m
- R = row spacing (in meters)
- Yield at each point is corrected to 12.5 % moisture [65,12,66], which is also the recommended harvest and storage moisture level by Grain South Africa [67] and the Agricultural Business Chamber of South Africa [68].
- Yield is adjusted by a factor of 0.95, which represents a conservative empirical bias adjustment applied to reduce potential overestimation associated with point-based objective yield surveys when scaled to field level. This factor acts as a reduction coefficient to account for field losses, harvesting inefficiencies, and non-uniformity in crop growth. This adjustment is consistent with the magnitude of estimation uncertainty reported by the CEC, for which final estimates are required to fall within ±5 % of reconciled end-of-season values [64].
- The 10,000 converts square meters (m^2) to hectares; 1000 converts grams to kilograms

$$\text{Field yield} = \frac{1}{5}\sum_{i=1}^{5} y_i \quad (2)$$

2.2.2. Sentinel-2 data processing in Google Earth Engine (GEE)

We used Sentinel-2 L2A surface reflectance (10–20 m) acquired through the GEE Platform [69]. The data covered 1 October to 30 June of the three cropping seasons (2018–2019, 2021–2022, and 2024–2025). To minimize cloud contamination, scenes with

CLOUDY_PIXEL_PERCENTAGE $\leq$ 70 % were retained, and the Scene Classification Layer (SCL) was used to mask clouds, cloud shadows, cirrus, and invalid pixels (SCL $\notin$ {0, 1, 3, 8, 9, 10, 11}). This reduced the risk of clouds contributing to the field-level means. For each field polygon (size range was 0.5 ha - 5 ha), we computed monthly means of seven indices by averaging valid field pixels. The indices include the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Green Normalized Difference Vegetation Index (GNDVI), Modified Chlorophyll Absorption Ratio Index (MCARI), Normalized Difference Red Edge Index (NDRE), Chlorophyll Index Red-Edge (CIrededge), and Moisture Stress Index (MSI). These indices capture different aspects of crop canopy structure, vigour, and water status (Table 1).

2.3. Interpolation and smoothing using python

Occasional missing monthly values were filled using linear interpolation. We then smoothed each field's monthly series using the Whittaker penalized least squares filter with a second-difference penalty and λ tuned in {50, 100, 500, 1000}. The smoothed and raw series across indices were visually compared, and λ = 500 was selected as the best compromise based on visual inspection across representative fields.

2.4. Computing seasonal metrics using python

Using each field's planting date, we restricted the metrics to post-planting months (month $\geq$ first day of planting month). This restriction ensured that phenological metrics reflected crop-related canopy dynamics rather than pre-season background vegetation or soil signals. All the computed phenological metrics are listed in Table 2.

2.5. Modelling methods

Model development was done in three stages to evaluate and improve the prediction of maize yield from Sentinel-2-derived phenological metrics:

(i) per-season modelling to capture intra-season yield variability
(ii) cross-season modelling to assess temporal generalization, and

Table 1
List of vegetation indices used for deriving phenological metrics.

Index and formulation	Description
$NDVI = \frac{B8 - B4}{B8 + B4}$	Most widely used vegetation index. It measures vegetation greenness and canopy vigour and is sensitive to chlorophyll content but saturates in dense vegetation [70].
$EVI = 2.5 \times \frac{B8 - B4}{B8 + 6 \times RB4 - 7.5 \times B2 + 1}$	Similar to NDVI with improved sensitivity in high biomass regions and corrects for atmospheric influences and soil background [71].
$GNDVI = \frac{B8 - B3}{B8 + B3}$	Similar to NDVI but uses the green band instead of the red. It is more sensitive to chlorophyll concentration and canopy stress [72].
$MCARI = [(B5 - B4) - 0.2 \times (B5 - B3)] \times \frac{B5}{B4}$	Designed to minimize the effect of non-photosynthetic background (soil, litter). Strongly related to chlorophyll content [73].
$RENDVI = \frac{B5 - B4}{B5 + B4}$	Uses red-edge bands (between red and NIR). More sensitive to canopy chlorophyll, reduces NDVI saturation [74].
$CIrededge = \frac{B8}{B5} - 1$	Strong proxy for leaf chlorophyll concentration. Often outperforms NDVI in dense canopies [75].
$MSI = \frac{B11}{B08}$	Simple ratio index, increases with vegetation water stress [76].

Bands; B2 = Blue, B3 = Green, B4 = Red, B5 = Red edge, B8 = Near Infrared (NIR), B11 = Short Wave Infrared (SWIR).

Table 2
List of phenological metrics that were derived from vegetation indices (VI).

Mathematical Formulation	Description
$MaxVI = \max_{t \in [t_0,\ t_n]} VI(t)$	Maximum VI value during the season (proxy for peak biomass/greenness).
$MeanVI = \frac{1}{n} \sum_{i=1}^{n} VI(t_i)$	Average VI over the growing season.
$SumVI = \int_{t_0}^{t_n} VI(t)dt \approx \sum_{i=1}^{n} VI(t_i)$	Cumulative VI across the growing season (related to productivity).
$StdVI = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (VI(t_i) - \overline{VI})^2}$	Standard deviation of VI values, indicating variability in the seasonal curve.
Start = t when VI_t > threshold	Start of season defined as the first date the VI exceeds 20 % of the seasonal amplitude above the minimum.
End = t when VI_t < threshold	End of season defined as the last date the VI falls below the same 20 % amplitude threshold during senescence
$TTP = t_{peak} - t_0$	Time to peak (TTP) in months from planting to maximum VI.
$Growth\ rate = \frac{MaxVI - StartVI}{t_{peak} - t_0}$	Rate of VI increase during green-up (slope of rise).
$Senescence\ rate = \frac{MaxVI - EndVI}{t_n - t_{peak}}$	Rate of VI decline during senescence.
$Late\ Season\ Mean = \frac{1}{k} \sum_{i=1}^{k} VI(t_i)$	Mean VI during early season (before peak).
$Early\ Season\ Mean = \frac{1}{m} \sum_{i=n-m+1}^{n} VI(t_i)$	Mean VI during late season (after peak).
$Amplitude = MaxVI - StartVI$	Difference between peak VI and baseline (seasonal intensity).
$NormAUC = \frac{\int_{t_0}^{t_n} VI(t)dt}{VI_{max} \bullet (t_n - t_0)}$	Area under the curve (AUC) of the VI normalized (Norm) by season length.
$t_{peak} = arg \max_{t \in [t_0, t_n]} VI(t)$	Calendar month when VI reaches maximum.
$PrePeak\ Mean = \frac{1}{t_{peak} - t_0} \int_{t_0}^{t_{peak}} VI(t)dt$	Mean VI before the seasonal peak.
$PostPeak\ Mean = \frac{1}{t_n - t_{peak}} \int_{t_{peak}}^{t_n} VI(t)dt$	Mean VI after the seasonal peak.
$PrePeakAUC = \int_{t_0}^{t_{peak}} VI(t)dt$	Cumulative VI before peak.
$PostPeakAUC = \int_{t_{peak}}^{t_n} VI(t)dt$	Cumulative VI after peak.

(iii) parsimonious modelling using the ten most influential predictors identified from the combined multi-season dataset.

All models were implemented in Python using the *scikit-learn* library [77].

2.5.1. Per-season modelling

Independent yield models were first trained for each of the three maize cropping season (2018–2019, 2021–2022, and 2024–2025). Since each dataset contained relatively few field observations (30–63), simpler and regularised linear models were preferred to prevent overfitting and maintain interpretability. Ridge (L2-regularised) and Least Absolute Shrinkage and Selection Operator (Lasso; L1-regularised) regressions were therefore applied [78–80]. Models were implemented within a preprocessing pipeline consisting of standardization, with median imputation included only as a precautionary safeguard to handle rare residual missing values arising from unresolved phenological timing metrics following interpolation and smoothing.

Model evaluation was primarily implemented using five-fold cross-validation (CV) to estimate out-of-fold (OOF) accuracy for each season.

To assess the robustness of performance estimates given the small sample sizes (n = 30–63), two additional validation strategies were applied: leave-one-out cross-validation (LOOCV), which maximizes the training set size per iteration and provides a complementary estimate of prediction error for small samples, and repeated five-fold CV (20 repetitions) to quantify the sensitivity of model performance to fold partitioning. These complementary validation schemes were used for diagnostic purposes and are reported in the Supplementary Material. These per-season models provided a baseline for understanding dominant phenological metric-yield relationships across farms within each growing season, reflecting variability in local management practices and field conditions. In this study, feature stability refers to the repeated high ranking of the same predictors across seasons, modelling approaches, and validation schemes. For each season, permutation importance was computed from the regularised models to rank predictors post hoc, producing season-specific importance patterns that were later compared to assess cross-season feature stability.

2.5.2. Cross-season modelling and feature selection

To quantify inter-annual consistency and model transferability, the three seasonal datasets were merged into a combined multi-season dataset (n = 123 fields). To enable cross-season modelling, phenological metrics were first restricted to those that were successfully computed in all three seasons. Metrics missing in any season were excluded, resulting in a common feature set that was consistently available across years. This intersection-based filtering ensured that all cross-season models were trained and evaluated using the same predictor variables, allowing fair comparison across modelling approaches. A suite of models of increasing complexity was tested, including Linear Regression, Ridge, Lasso, Elastic Net, Random Forest (RF), and Gradient Boosting (GB) regression. Model performance was assessed using standard five-fold cross-validation and leave-one-season-out (LOSO) cross-validation, with LOSO serving as the primary framework for evaluating temporal transferability.

Unlike the per-season analyses, feature importance in the combined multi-season setting was explicitly evaluated to identify predictors that were both informative and stable across years. Two complementary approaches were applied: permutation importance [81] and SHapley Additive exPlanations (SHAP [82]). Permutation importance quantified the marginal contribution of each predictor by measuring performance degradation following feature permutation, while SHAP analysis, conducted using the GB model, captured nonlinear effects and interactions and provided a global ranking based on mean absolute SHAP values. Concordance between permutation importance and SHAP rankings was assessed qualitatively to identify predictors that were consistently influential across seasons and modelling approaches. The ten most influential predictors identified from this combined multi-season analysis were selected for subsequent parsimonious modelling.

2.5.3. Parsimonious modelling using stable cross-season predictors

Based on the global feature-importance analysis described in Section 2.5.2, a parsimonious modelling experiment was conducted in which only the ten most influential predictors identified from the combined multi-season dataset were retained as fixed inputs. These predictors were treated as a reduced, stable feature set for evaluating whether compact models could retain predictive performance under inter-annual variability. The principle of parsimony (or Occam's Razor) states that, among competing models, the simplest model that adequately explains the data, should be preferred [83]. Accordingly, GB and Ridge regression were retrained using only the Top-10 predictors to assess whether a compact and interpretable feature set could achieve performance comparable to models trained on the full set of seasonal metrics, without additional feature selection. Model performance was evaluated using five-fold cross-validation with out-of-fold predictions to assess overall robustness. In addition, LOSO cross-validation was applied as the primary evaluation of temporal transferability.

2.6. Cross-validation and evaluation metrics

For each modelling stage, performance was summarized using coefficient of determination (R^2): proportion of variance explained by the model [84], root mean square error (RMSE): average magnitude of prediction error [85], and mean absolute error (MAE): absolute deviation from observed yield [85]. Final R^2, RMSE, and MAE values were calculated from pooled out-of-fold predictions rather than by averaging fold-wise metrics, ensuring unbiased performance estimation. Scatterplots of observed VS. predicted yields and residual distributions were also examined to assess bias and stability across the yield range.

3. Results

Observed maize yields across the three seasons ranged from 367 to 9080 kg ha^{-1}, with seasonal mean yields increasing from 3120 kg ha^{-1} in 2018–2019 to 4735 kg ha^{-1} in 2024–2025 (Table 3). The average yield of the all the 123 farms was 4211.70 kg ha^{-1}.

3.1. Per-season models

Regularised regression achieved consistent yield prediction accuracy across all three seasons (Table 4). Ridge regression yielded the best results for 2018–2019 (R^2 = 0.56, RMSE = 1148 kg ha^{-1}), while Lasso regression performed best for 2021–2022 (R^2 = 0.53, RMSE = 1263 kg ha^{-1}) and 2024–2025 (R^2 = 0.57, RMSE = 1329 kg ha^{-1}).

Predicted yields closely followed the 1:1 line, indicating that regularised models effectively captured dominant statistical relationships between seasonal canopy dynamics and observed yield (Fig. 2). However, a tendency toward underestimation at the upper end of the yield range was observed, a pattern commonly associated with regularization and limited sample sizes. Because model evaluation in Fig. 2 is based on cross-validation and relatively small seasonal sample sizes, performance estimates may be subject to optimistic bias. To assess the robustness of these results, additional validation using leave-one-out cross-validation (LOOCV) and repeated five-fold cross-validation was conducted. These supplementary analyses yielded performance estimates consistent with the primary five-fold CV results (Supplementary Tables S1-S2), indicating that the reported accuracies are not driven by favourable fold partitioning. All per-season performance metrics and predicted-observed relationships shown in Fig. 2 are based on out-of-fold predictions from cross-validation rather than independent test sets.

Across all three seasons, phenological metrics describing seasonal intensity (area-under-the-curve-type metrics) and pre- or post-peak canopy dynamics consistently ranked among the most influential predictors of maize yield (Fig. 3). Permutation importance analysis was used to rank phenological metrics within each season, providing an interpretive view of season-specific predictor contributions rather than a basis for feature selection. In the 2018–2019 season, MSI and NDVI-derived metrics ranked highly. In 2021–2022, MCARI-based metrics emerged as dominant predictors. In 2024–2025, CIrededge-based metrics ranked highest, indicating an increased model sensitivity to red-edge spectral information. When importance scores were averaged across seasons, red-edge-derived indices (particularly CIrededge and MCARI) consistently exhibited high rankings, demonstrating consistent importance across seasons (Fig. 3).

Table 3
Observed yields by season.

Season	Minimum (kg ha^{-1})	Maximum (kg ha^{-1})	Average (kg ha^{-1})
2018–2019	367.11	6558.88	3119.84
2021–2022	871.15	7836.81	4220.47
2024–2025	444.47	9080.41	4734.92

Table 4
Seasonal maize yield predictions.

Season	Best model	R^2	RMSE (kg ha^{-1})	MAE (kg ha^{-1})	Farms
2018–2019	Ridge	0.56	1148	880	30
2021–2022	Lasso	0.53	1263	1023	30
2024–2025	Lasso	0.57	1329	1125	63

3.2. Cross-season models

When all three seasons were merged into a single dataset ($n = 123$ fields), both linear and ensemble models (represented here by GB regression) produced comparable accuracies under five-fold CV (Table 5). The best performance was achieved by the Elastic Net model ($R^2 = 0.52$, RMSE = 1401 kg ha^{-1}), followed closely by Lasso ($R^2 = 0.51$, RMSE = 1426 kg ha^{-1}) and GB ($R^2 = 0.45$, RMSE = 1447 kg ha^{-1}).

3.3. Feature stability and parsimonious models

To investigate whether a reduced set of predictors could retain explanatory relevance under inter-annual variability, a pooled analysis was conducted using observations from all three seasons. Feature-importance results derived from this multi-season dataset are summarized in Fig. 4. Using a GB model fitted to the combined dataset, feature contributions were quantified using SHAP values and complemented by permutation importance. Features that remained highly ranked under this pooled setting were interpreted as exhibiting cross-season consistency in their contribution to yield variation, rather than importance

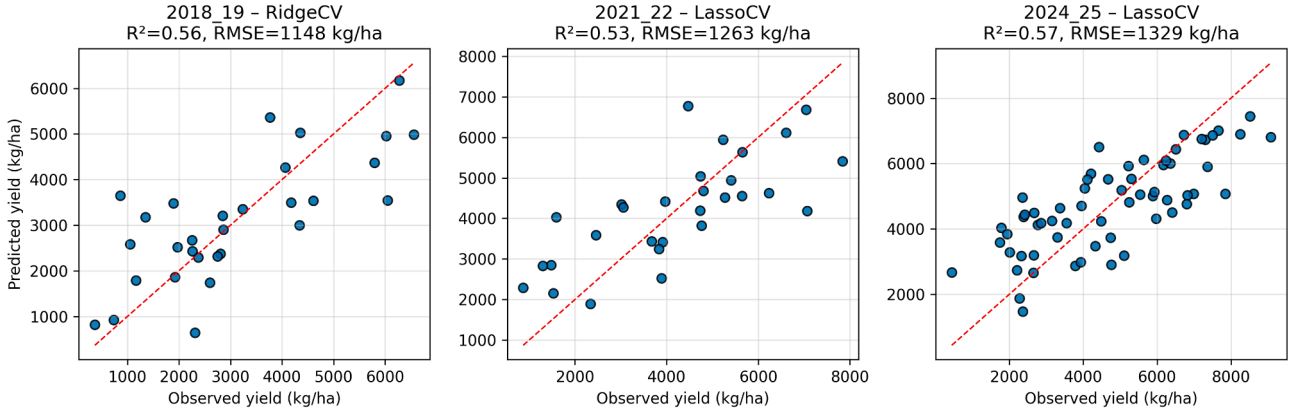

Fig. 2. Seasonal maize yield predictions from per-season models. The three panels show observed versus predicted yield (kg ha^{-1}) for each growing season. Predictions are out-of-fold (OOF) estimates from five-fold cross-validation (CV). The 1:1 line indicates perfect agreement.

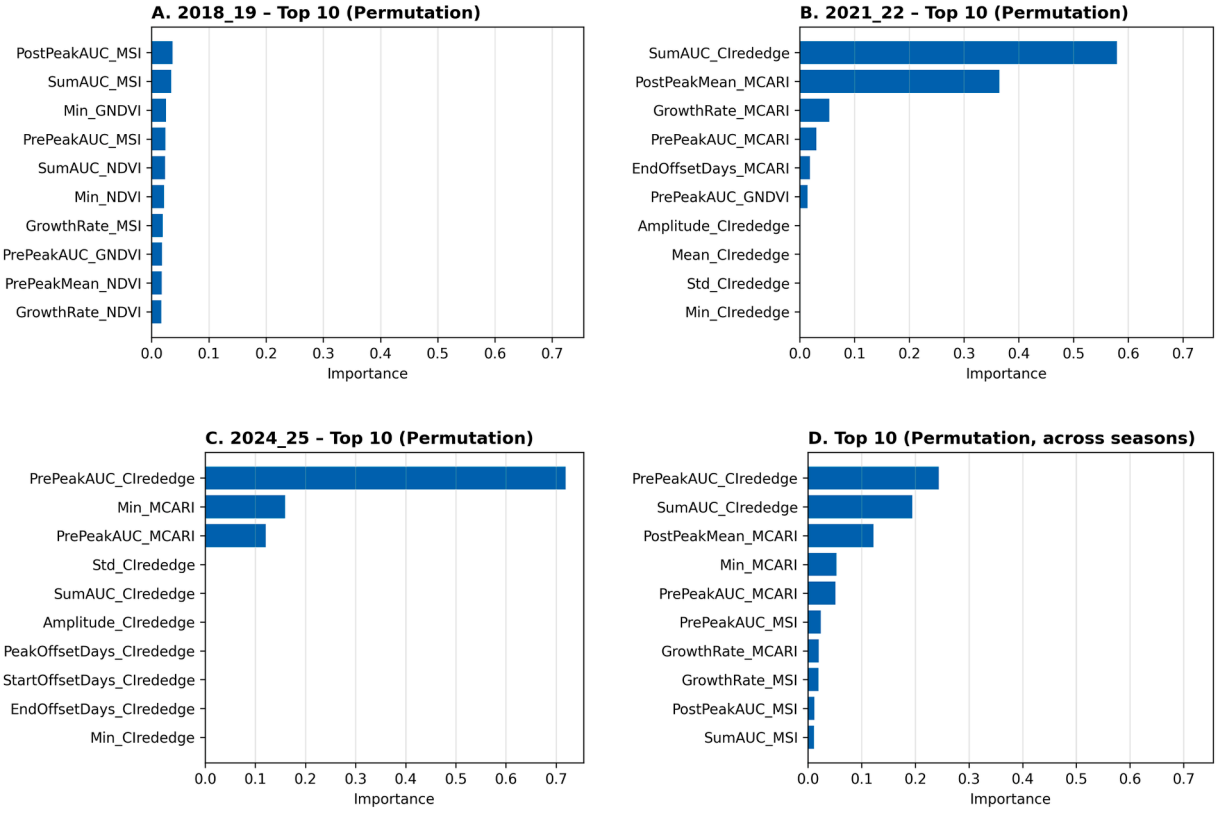

Fig. 3. Top features contributing to maize yield predictions per season (Features with near-zero permutation importance are retained to illustrate sparsity and seasonal selectivity).

Table 5
Combined-season maize yield prediction models.

Model	CV Type	R^2	RMSE	MAE
ElasticNet	5Fold	0.52	1401	1141
Lasso	5Fold	0.51	1426	1130
GB	5Fold	0.49	1447	1192
RF	5Fold	0.49	1484	1189
Ridge	5Fold	0.46	1563	1190
ElasticNet	LOSO	0.41	1639	1262
Lasso	LOSO	0.35	1639	1302
GB	LOSO	0.30	1699	1350
RF	LOSO	0.28	1714	1393
Ridge	LOSO	0.18	1836	1462

restricted to a single growing season. The ten most influential predictors identified in this analysis are shown in Fig. 4.

To assess whether a compact predictor set could maintain predictive performance comparable to full-feature models, Ridge and GB regressions were subsequently retrained using only these Top-10 features and evaluated using five-fold cross-validation. The parsimonious Ridge model achieved accuracy ($R^2 = 0.56$, RMSE = 1343 kg ha^{-1}) comparable to the best-performing per-season models and outperforming the full-feature GB and Random Forest models evaluated on the combined dataset (Table 6).

Out-of-fold predictions closely followed the 1:1 line, with minimal bias across the yield range, demonstrating that most explanatory power is retained within a small subset of spectrally interpretable variables (Fig. 5). When compared to the full-feature GB baseline, the reduction to only ten predictors resulted in comparable predictive accuracy.

3.4. *Temporal transferability (Leave-One-Season-Out validation)*

Both Ridge Top-10 and GB Top-10 models showed limited and season-dependent transferability across seasons. Predictive accuracy varied by test year, with differences observed across seasons. When the 2018–2019 season was held out, model performance decreased (Table 7). Testing on the 2021–2022 season improved performance for Ridge, while the 2024–2025 season achieved low to moderate generalization (Table 7). Overall, the mean LOSO accuracy across seasons was $R^2 = 0.32$ for GB and $R^2 = 0.46$ for Ridge (Table 7; Fig. 6).

3.5. *Residual analysis*

The residual distributions of the Ridge Top-10 and GB Top-10 models were centred near zero, indicating minimal systematic bias in yield predictions (Fig. 7). Mean residuals were −38 kg ha^{-1} for Ridge and −42 kg ha^{-1} for GB, reflecting slight average underestimation. The dispersion of residuals was moderate, with standard deviations of 1348 kg ha^{-1} and 1448 kg ha^{-1} for Ridge and GB, respectively, suggesting similar variability in predictive performance. However, residual dispersion remained substantial, with standard deviations of 1348 kg ha^{-1} and 1448 kg ha^{-1}, respectively, highlighting considerable unexplained variability at the field level. Residuals ranged from −3205 to +2589 kg ha^{-1} for Ridge and from −4521 to +2755 kg ha^{-1} for GB, indicating the presence of a small number of extreme errors. These patterns suggest that while predictions are not systematically biased, model uncertainty remains high for individual fields, particularly at the upper end of the yield distribution.

4. Discussion

This study demonstrates the potential use of Sentinel-2-derived, systematically selected seasonal metrics in the estimation of smallholder maize yields, while highlighting important limitations related to data availability and inter-annual variability. By leveraging time-series vegetation indices processed through GEE and transformed into phenological metrics, the study achieved moderate predictive accuracy, with seasonal models yielding R^2 values of 0.53 to 0.57 and RMSEs of 1148 to 1329 kg ha^{-1}. This finding indicates that a substantial proportion of yield variability among smallholder fields can be statistically associated with regularised linear models trained on spectrally derived, systematically selected, and physiologically interpretable features. The alignment of predicted yields with the 1:1 line in the scatterplots suggests that the models captured a meaningful component of intra-seasonal yield variability, although a substantial proportion of variance remained unexplained.

Table 6
Performance of parsimonious Top-10 models evaluated using five-fold cross-validation. The Top-10 predictors were selected based on their consistent importance across seasons, as identified using Gradient Boosting (GB) feature-importance analysis interpreted with SHAP values and complemented by permutation importance.

Model	R^2	RMSE (kg ha^{-1})	MAE (kg ha^{-1})
Ridge Top-10	0.56	1343	1057
GB Top-10	0.49	1442	1175
GB (All features)	0.49	1447	1192

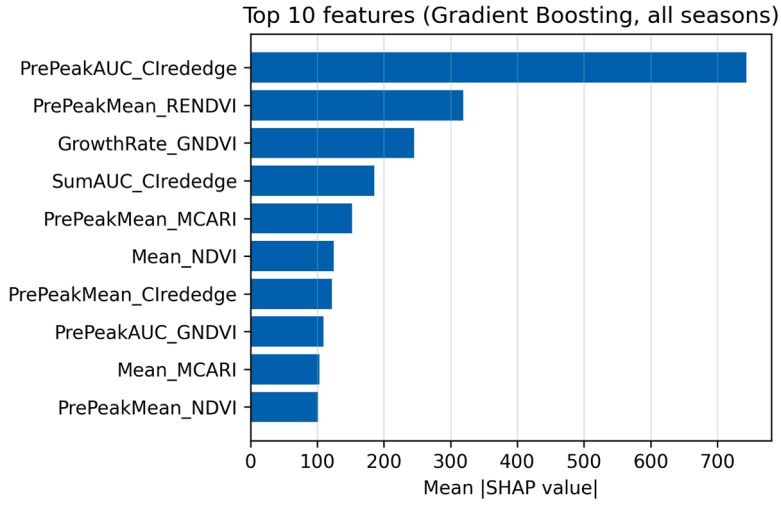

Fig. 4. Top-10 predictors of maize yield identified from a GB model trained on the merged multi-season dataset ($n = 123$), ranked by mean absolute SHAP values and confirmed using permutation importance.

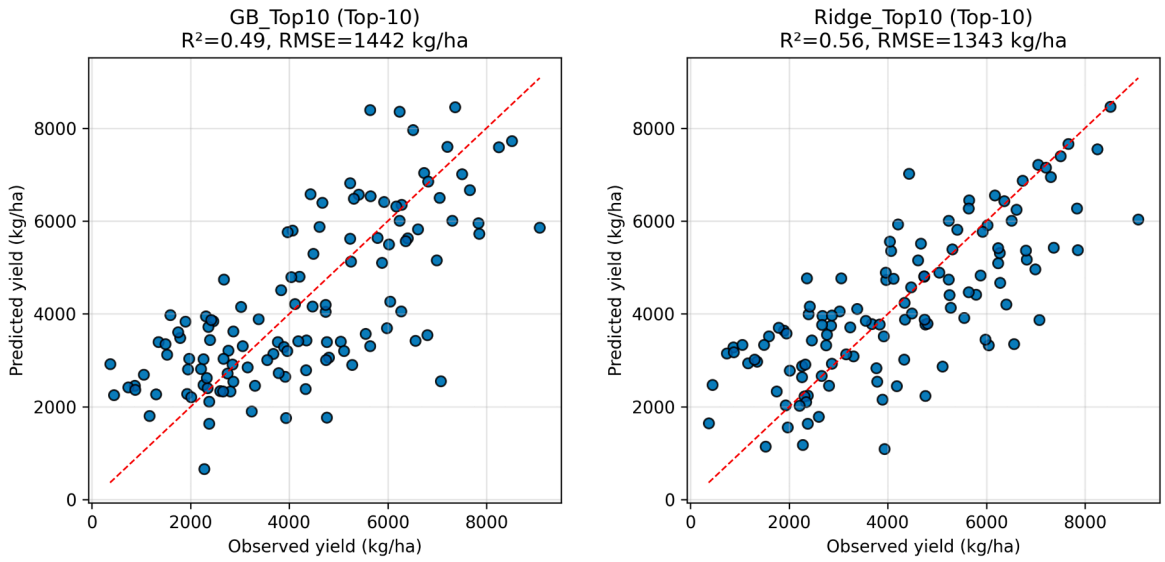

Fig. 5. Performance of parsimonious models using the Top-10 predictors. Observed versus predicted yield (kg ha^{-1}) for the parsimonious Ridge Top-10 (left) and Gradient Boosting (GB, right) Top-10, evaluated using five-fold cross-validation (CV) with out-of-fold (OOF) predictions. The 1:1 line indicates perfect agreement.

Table 7
Summary of leave-one-season-out (LOSO) model performance per test season. Gradient Boosting (GB) and Ridge trained with Top-10 features.

Model	Test season	R^2 LOSO	RMSE	MAE	Fields
GB Top10	2018–2019	0.34	1410	1151	30
Ridge Top10	2018–2019	0.30	1454	1215	30
GB Top10	2021–2022	0.34	1510	1228	30
Ridge Top10	2021–2022	0.63	1130	866	30
GB Top10	2024–2025	0.27	1734	1416	63
Ridge Top10	2024–2025	0.44	1511	1185	63

Across all three seasons, permutation importance and SHAP analyses consistently ranked features describing pre-peak growth and season-long cumulative vegetation signals among the most influential predictors. At the seasonal level, phenological metrics derived from MSI and NDVI were most influential in 2018–2019, MCARI in 2021–2022, and CIrededge in 2024–2025. There are plausible explanations for these differences. Firstly, this pattern is consistent with the idea that different seasons foreground different limiting factors [86,87], so the vegetation index family that best tracks the dominant constraint becomes most predictive. For example, seasons in which water availability and canopy moisture dynamics are more influential may elevate MSI and NDVI [76,88,89]. When variability in chlorophyll absorption and canopy structure better differentiates field performance, indices such as MCARI may become more informative [73]. The stronger appearance of red-edge-based metrics (e.g., CIrededge integrals) in another season can similarly reflect conditions where chlorophyll-related spectral variation is clearer in the Sentinel-2 signal. Importantly, these shifts should be interpreted as season-specific statistical associations rather than direct evidence of causal physiological control, and they may also be amplified by practical factors such as cloud-related data gaps, planting-date spread, and within-season management heterogeneity that affect the clarity of phenological trajectories.

When feature importance was averaged across seasons, red-edge-derived indices (CIrededge and MCARI) consistently ranked highest, followed by GNDVI and MSI metrics. Across models, PrePeakAUC_CIrededge, SumAUC_CIrededge, PrePeakAUC_MCARI, Min_MCARI and PostPeak_MCARI emerged as the most stable predictors of yield, capturing pre-peak accumulation and season-long canopy productivity. The recurrence of these metrics across seasons and algorithms suggests that yield variation was most strongly associated with spectral patterns capturing early-season canopy development and cumulative seasonal vegetation activity. Complementary SHAP rankings from the combined multi-season GB analysis were broadly consistent with the permutation-based patterns, with PrePeakAUC_CIrededge, PrePeakMean_RENDVI, and GrowthRate_GNDVI among the top contributors. Taken together, these results indicate that integrative pre-peak and season-long metrics provided a compact and interpretable predictor set for yield modelling, although the LOSO results show that cross-season generalization remains season-dependent.

When merging data across seasons, cross-season models achieved comparable predictive accuracy, with Elastic Net performing best (R^2 = 0.52, RMSE = 1401 kg ha^{-1}). GB showed slightly lower performance (R^2 = 0.45), indicating that added model flexibility did not translate into improved generalization in this merged setting. Parsimonious models trained on the ten highest-ranked predictors retained comparable accuracy (R^2 = 0.49 to 0.56, RMSE = 1343 to 1442 kg ha^{-1}), supporting the conclusion that a compact subset of phenological metrics can capture much of the yield-predictive signal present in the data. This parsimony improves interpretability and reduces computational demands; however, given the observed residual dispersion and season-dependent transferability, these results are best interpreted as evidence of potential utility rather than fully operational readiness without additional seasons, covariates, and independent evaluation.

An important consideration in this framework is the use of field-reported planting dates to align vegetation-index time series and restrict phenological metric extraction to post-planting periods. Although planting date was not included as a predictor in the yield models, it represents ancillary field information that is not directly observed by Sentinel-2. Its use helps ensure that extracted metrics reflect crop canopy development rather than pre-season background signals. However, this reliance introduces an external input that may limit fully automated implementation where planting dates are unavailable. Future work could reduce this dependency by incorporating satellite-based planting date detection or phenology-based season onset estimation.

Testing temporal transferability using LOSO validation revealed strong season-dependent performance, with Ridge regression outperforming GB in some test seasons but not consistently across years. This variability indicates that yield-phenology relationships in smallholder systems can shift across seasons, and that model generalization is constrained by inter-annual differences in yield distributions, phenological signal clarity, and unmeasured sources of heterogeneity (e.g.,

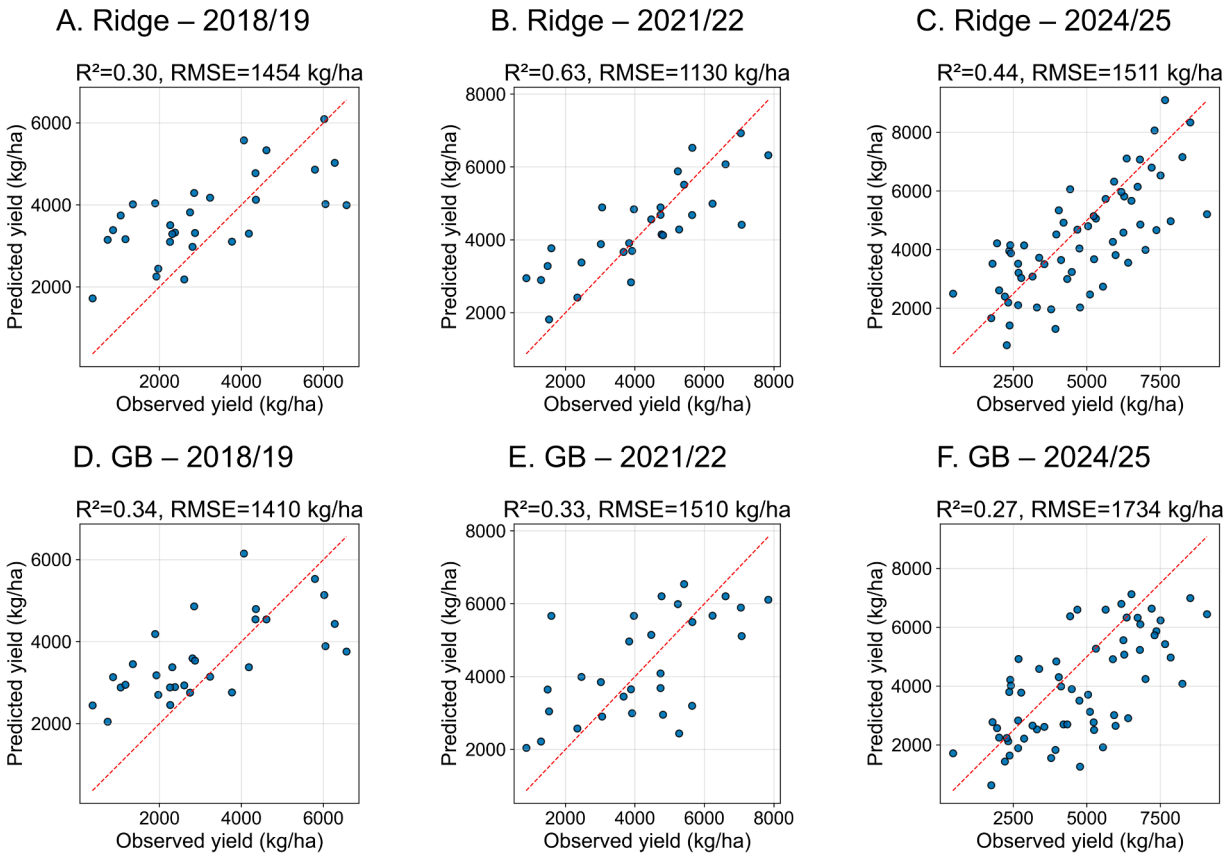

Fig. 6. Leave-one-season-out (LOSO) validation results by test season. Panels show observed versus predicted yield (kg ha^{-1}) when models are trained on two seasons and evaluated on the held-out season (test year shown in each panel). Ridge Top-10 and Gradient Boosting (GB) Top-10 results are displayed. The 1:1 line indicates perfect agreement.

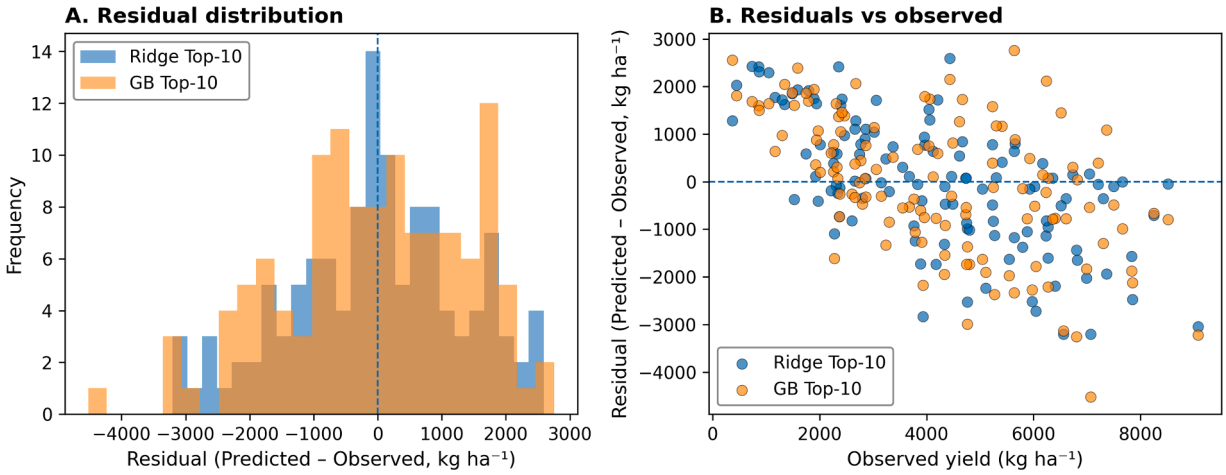

Fig. 7. Residual diagnostics for parsimonious Ridge and *Gradient Boosting (GB) models.*

management and weather variability) that are not captured by vegetation indices alone. Residual diagnostics indicated limited systematic bias (mean residuals close to zero) but substantial dispersion, with a small number of large errors, highlighting considerable unexplained field-level variability. Overall, these findings support cautious interpretation of cross-season generalization and reinforce the importance of season-aware validation when deploying phenology-based yield models.

Overall, the results show that Sentinel-2-derived phenological metrics, combined with regularised linear models and interpretable feature analyses, can provide useful yield estimates in fragmented smallholder landscapes where field data are limited. Relative to coarser-resolution sensors, Sentinel-2 offers spatial detail that is better aligned with small field sizes, supporting field-level variation in crop trajectories. The integration of a mobile application for field data collection improves the traceability and standardisation of ground reference yields, strengthening the linkage between remote sensing predictors and observed outcomes. Nevertheless, the residual dispersion and LOSO variability indicate that vegetation indices alone do not capture all drivers of yield variability. Incorporating additional covariates (e.g., weather, soils, management proxies) and expanding multi-year field sampling would

likely reduce unexplained variance and improve transferability, supporting more reliable use in operational monitoring contexts.

5. Conclusion

This study shows that Sentinel-2–derived phenological metrics, processed in GEE and analysed using regularised linear and ensemble models, can provide interpretable, satellite-based estimates of smallholder maize yield variability in South Africa. Per-season modelling achieved R^2 values of 0.53–0.57, and pooled cross-season modelling reached $R^2 = 0.52$ with Elastic Net under five-fold cross-validation. Across seasons, the most consistently influential predictors were pre-peak and season-integrated CIrededge and MCARI metrics, indicating that red-edge and chlorophyll-sensitive signals linked to canopy development were repeatedly associated with yield differences among fields. Parsimonious Top-10 models retained comparable accuracy, while LOSO evaluation showed strong season dependence in transferability; residuals were generally centred near zero (limited systematic bias) but remained widely dispersed, indicating substantial unexplained field-level variability.

Overall, the results indicate the potential of Sentinel-2 phenological metrics as a cost-effective complement to labour-intensive surveys and complex process-based models for smallholder settings. While further gains in operational reliability will likely require larger multi-year samples and additional covariates (e.g., weather, soils, and management), the framework provides a practical basis for strengthening yield monitoring in under-documented smallholder systems.

5.1. Limitations and future work

While the results are promising, several limitations should be noted. First, the analysis covered only three seasons ($n = 123$ fields), which limits statistical generalisation and constrains the ability to characterise the full range of inter-annual variability observed in smallholder systems. Second, predictors were derived exclusively from optical Sentinel-2 vegetation indices, which can be affected by cloud contamination and may saturate under dense canopies; integrating complementary sensors (e.g., Sentinel-1 and/or thermal products where available) may improve robustness under variable observation conditions. Third, important covariates such as soils, weather, and management practices were not explicitly included in the models and likely explain a substantial portion of the residual dispersion and season-dependent transferability observed under LOSO evaluation. Fourth, field-level averages were modelled, which can mask within-field heterogeneity; where appropriate and feasible, higher-resolution observations could support sub-field assessment and improved calibration, although such data sources may involve cost and continuity trade-offs. Finally, the current modelling framework prioritised interpretability using regularised linear and ensemble methods; future studies with larger multi-year datasets could explore hybrid or more flexible approaches while maintaining rigorous cross-season evaluation.

Data and code availability

The Python code used for data processing, modelling, and feature-importance analysis in this study is publicly available at: https://github.com/masizawonga63-eng/sentinel2-maize-yield-phenology. Due to farmer confidentiality agreements and ethical restrictions, the raw field-level yield data and associated farm identifiers cannot be publicly shared. Derived phenological metrics and aggregated results supporting the findings of this study are available from the corresponding author upon reasonable request.

Funding

This work was supported by the Department of Agriculture's Policy Research and Analysis directorate (Project number: A255).

Ethics in publishing statement

I testify on behalf of all co-authors that our article submitted followed ethical principles in publishing.

Title: Predicting Smallholder Maize Yield Using Sentinel-2-derived Phenological Metrics

All authors agree that:

This research presents an accurate account of the work performed, all data presented are accurate and methodologies detailed enough to permit others to replicate the work.

This manuscript represents entirely original works and or if work and/or words of others have been used, that this has been appropriately cited or quoted and permission has been obtained where necessary.

This material has not been published in whole or in part elsewhere.

The manuscript is not currently being considered for publication in another journal.

That generative AI and AI-assisted technologies have not been utilized in the writing process or if used, disclosed in the manuscript the use of AI and AI-assisted technologies and a statement will appear in the published work.

That generative AI and AI-assisted technologies have not been used to create or alter images unless specifically used as part of the research design where such use must be described in a reproducible manner in the methods section.

All authors have been personally and actively involved in substantive work leading to the manuscript and will hold themselves jointly and individually responsible for its content.

Date: 04 November 2025

CRediT authorship contribution statement

Wonga Masiza: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Basani Lammy Nkuna:** Writing – review & editing, Writing – original draft, Validation, Data curation. **Phathutshedzo Eugene Ratshiedana:** Writing – review & editing, Validation, Investigation, Data curation. **Akhona Madasa:** Writing – review & editing, Data curation. **Lwandile Nduku:** Writing – review & editing, Investigation, Data curation. **Tumelo Shwatja:** Writing – review & editing, Investigation, Data curation. **Johannes George Chirima:** Writing – review & editing, Validation, Project administration, Conceptualization. **Adolph Nyamugama:** Writing – review & editing. **Khaled Abutaleb:** Writing – review & editing, Validation, Investigation, Funding acquisition, Conceptualization. **Pitso Walter Khoboko:** Writing – review & editing, Software, Funding acquisition, Data curation, Conceptualization. **Hamisai Hamandawana:** Writing – review & editing, Visualization, Validation, Investigation.

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.

Acknowledgements

The authors wish to thank South Africa's Agricultural Research Council for supporting this work, and the Department of Agriculture's Policy Research and Analysis Directorate for funding the research activities that enabled the writing of this paper. We are grateful to Mr Sibongiseni Ndimande, Ms Khathutshelo Patience Mphumbude and Ms Makgomo Vivian Mosoma for project coordination and administration.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2026.101870.

Data availability

Data will be made available on request.

References

[1] UN. (2025). The sustainable development goals report 2025. https://sdgs.un.org/goals.

[2] UN. (2015). The 2030 agenda for sustainable development's 17 sustainable development goals (SDGs). https://sdgs.un.org.

[3] Fenta, A.A., Tsunekawa, A., Haregeweyn, N., Tsubo, M., Yasuda, H., Shimizu, K., Kawai, T., Ebabu, K., Berihun, M.L., Sultan, D., Belay, A.S., & Sun, J. (2020). Cropland expansion outweighs the monetary effect of declining natural vegetation on ecosystem services in sub-Saharan Africa. *Ecosystem Services*, 45. https://doi.org/10.1016/j.ecoser.2020.101154.

[4] A. Molotoks, E. Stehfest, J. Doelman, F. Albanito, N. Fitton, T.P. Dawson, P. Smith, Global projections of future cropland expansion to 2050 and direct impacts on biodiversity and carbon storage, Glob. Chang. Biol. 24 (12) (2018) 5895–5908, https://doi.org/10.1111/gcb.14459.

[5] W. Wu, Q. Yu, L. You, K. Chen, H. Tang, J. Liu, Global cropping intensity gaps: increasing food production without cropland expansion, Land. Use Policy 76 (2018) 515–525, https://doi.org/10.1016/j.landusepol.2018.02.032.

[6] C. Carletto, Better data, higher impact: improving agricultural data systems for societal change, Eur. Rev. Agric. Econ. 48 (4) (2021) 719–740, https://doi.org/10.1093/erae/jbab030.

[7] Carletto, C., Jolliffe, D., & Banerjee, R. (2013). The Emperor has no data! agricultural statistics in Sub-Saharan Africa. In *World Bank Working Paper*. ht tps://doi.org/Available online: http://mortenjerven.com/wp-content/uploads/2013/04/Panel-3-Carletto.pdf (Accessed on 16 May 2025).

[8] C. Carletto, D. Jolliffe, R. Banerjee, From tragedy to renaissance: improving agricultural data for better policies, J. Dev. Stud. 51 (2) (2015) 133–148, https://doi.org/10.1080/00220388.2014.968140.

[9] E.C. Geyman, A. Ferris, R. Sahajpal, W. Anderson, D. Lee, N. Hausmann, An Africa-wide agricultural production database to support policy and satellite-based measurement systems, Sci. Data 12 (1) (2025) 1–22, https://doi.org/10.1038/s41597-025-05257-5.

[10] L. Tandzi, C. Mutengwa, Estimation of maize (Zea mays L.) yield per harvest area: appropriate methods. Agronomy, MDPI AG, 2020, https://doi.org/10.3390/agronomy10010029. Vol. 10, Number 1.

[11] F. Kosmowski, J. Chamberlin, H. Ayalew, T. Sida, K. Abay, P. Craufurd, How accurate are yield estimates from crop cuts? evidence from smallholder maize farms in Ethiopia, Food Policy (2021) 102, https://doi.org/10.1016/j.foodpol.2021.102122.

[12] K. Tiedeman, J. Chamberlin, F. Kosmowski, H. Ayalew, T. Sida, R.J. Hijmans, Field data collection methods strongly affect satellite-based crop yield estimation, Remote Sens. (Basel) 14 (9) (2022) 1–17, https://doi.org/10.3390/rs14091995.

[13] Y.L. Gavasso-Rita, S.M. Papalexiou, Y. Li, A. Elshorbagy, Z. Li, C. Schuster-Wallace, Crop models and their use in assessing crop production and food security: a review. Food and Energy Security, John Wiley and Sons Inc, 2024, https://doi.org/10.1002/fes3.503. Vol. 13, Number 1.

[14] B. Maestrini, G. Mimić, P.A.J. van Oort, K. Jindo, S. Brdar, F.K. van Evert, I. Athanasiados, Mixing process-based and data-driven approaches in yield prediction. European Journal of Agronomy, Elsevier B.V., 2022, https://doi.org/10.1016/j.eja.2022.126569. Vol. 139.

[15] M.A. Jabed, M.A. Azmi Murad, Crop yield prediction in agriculture: a comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. Heliyon, Elsevier Ltd, 2024, https://doi.org/10.1016/j.heliyon.2024.e40836. Vol. 10, Number 24.

[16] T. van Klompenburg, A. Kassahun, C. Catal, Crop yield prediction using machine learning: a systematic literature review, Comput. Electron. Agric. 177 (August) (2020) 105709, https://doi.org/10.1016/j.compag.2020.105709.

[17] A.M. Ali, M. Abouelghar, A.A. Belal, N. Saleh, M. Yones, A.I. Selim, M.E.S. Amin, A. Elwesemy, D.E. Kucher, S. Maginan, I. Savin, Crop yield prediction using multi sensors remote sensing (review article), Egypt. J. Remote Sens. Space Sci. 25 (3) (2022) 711–716, https://doi.org/10.1016/j.ejrs.2022.04.006.

[18] P. Muruganantham, S. Wibowo, S. Grandhi, N.H. Samrat, N. Islam, A systematic literature review on crop yield prediction with deep learning and remote sensing. Remote Sensing, MDPI, 2022, https://doi.org/10.3390/rs14091990. Vol. 14, Number 9.

[19] Sapkota, T.B., Jat, M.L., Jat, R.K., Kapoor, P., & Stirling, C. (2016). Yield estimation of food and non-food crops in smallholder production systems. In *Methods For Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture* (pp. 163–174). https://doi.org/10.1007/978-3-319-29794-1_8.

[20] P.N. Kephe, K.K. Ayisi, B.M. Petja, Challenges and opportunities in crop simulation modelling under seasonal and projected climate change scenarios for crop production in South Africa, Agric. Food Secur. 10 (1) (2021) 1–24, https://doi.org/10.1186/s40066-020-00283-5.

[21] D.B. Lobell, G. Azzari, M. Burke, S. Gourlay, Z. Jin, T. Kilic, S. Murray, Eyes in the sky, boots on the ground: assessing satellite- and ground-based approaches to crop yield measurement and analysis, Am. J. Agric. Econ. 102 (1) (2020) 202–219, https://doi.org/10.1093/ajae/aaz051.

[22] H. Burdett, C. Wellen, Statistical and machine learning methods for crop yield prediction in the context of precision agriculture, Precis. Agric. 23 (5) (2022) 1553–1574, https://doi.org/10.1007/s11119-022-09897-0.

[23] T. Hu, X. Zhang, S. Khanal, R. Wilson, G. Leng, E.M. Toman, X. Wang, Y. Li, K. Zhao, Climate change impacts on crop yields: a review of empirical findings, statistical crop models, and machine learning methods. Environmental Modelling and Software, Elsevier Ltd, 2024, https://doi.org/10.1016/j.envsoft.2024.106119. Vol. 179.

[24] L. Karthikeyan, I. Chawla, A.K. Mishra, A review of remote sensing applications in agriculture for food security: crop growth and yield, irrigation, and crop losses, J Hydrol 586 (March) (2020) 124905, https://doi.org/10.1016/j.jhydrol.2020.124905.

[25] Q. Dai, H. Chen, Z. Chen, C. Liu, G. Li, Y. Wang, X. Hu, Identification of optimal phenological periods for summer maize yield prediction using UAV-based multispectral data, J. Integr. Agric. (2025), https://doi.org/10.1016/j.jia.2025.02.026.

[26] Z. Ji, Y. Pan, X. Zhu, J. Wang, Q. Li, Prediction of crop yield using phenological information extracted from remote sensing vegetation index, Sens. (Switz.) 21 (4) (2021) 1–17, https://doi.org/10.3390/s21041406.

[27] F. Kuri, A. Murwira, K.S. Murwira, M. Masocha, Accounting for phenology in maize yield prediction using remotely sensed dry dekads, Geocarto Int. 33 (7) (2018) 723–736, https://doi.org/10.1080/10106049.2017.1299798.

[28] Pei, J., Tan, S., Zou, Y., Liao, C., He, Y., Wang, J., Huang, H., Wang, T., Tian, H., Fang, H., Wang, L., & Huang, J. (2025). The role of phenology in crop yield prediction: comparison of ground-based phenology and remotely sensed phenology. *Agricultural and Forest Meteorology*, 361. https://doi.org/10.1016/j.agrformet.2024.110340.

[29] U. Meier, H. Bleiholder, L. Buhr, C. Feller, H. Hack, M. Heß, P.D. Lancashire, U. Schnock, R. Stauß, T. van den Boom, E. Weber, P. Zwerger, C. Peter Zwerger, The BBCH system to coding the phenological growth stages of plants-history and publications, J. FÜR KULT. (2) (2009) 41–52, https://doi.org/10.5073/JfK.2009.02.01.

[30] H. Zhao, F. Gao, M. Anderson, R. Cirone, G.J. Chang, Improving crop condition monitoring using phenologically aligned vegetation index anomalies – A case study in central Iowa, Int. J. Appl. Earth Obs. Geoinf. (2025) 139, https://doi.org/10.1016/j.jag.2025.104526.

[31] Albarenque, S., Basso, B., Davidson, O., Maestrini, B., & Melchiori, R. (2023). Plant emergence and maize (Zea mays L.) yield across multiple farmers' fields. *Field Crops Research*, 302. https://doi.org/10.1016/j.fcr.2023.109090.

[32] H. Li, L. Shao, X. Liu, H. Sun, S. Chen, X. Zhang, What matters more, biomass accumulation or allocation, in yield and water productivity improvement for winter wheat during the past two decades», Eur. J. Agron. (2023) 149, https://doi.org/10.1016/j.eja.2023.126910.

[33] X. Ma, J. Jin, X. Zhu, Y. Zhou, Q. Xie, Remote sensing of land surface phenology: editorial, Remote Sens. (Basel) 14 (17) (2022), https://doi.org/10.3390/rs14174310.

[34] Meier, G.A., & Brown, J.F. (2014). *Remote sensing of land surface phenology* (Fact Sheet). https://doi.org/10.3133/fs20143052.

[35] Zhao, Q., Qu, Y., & Liu, D. (2025). Real-time monitoring of maize phenology using ground camera fusion information. *Smart Agricultural Technology*, 10. https://doi.org/10.1016/j.atech.2025.100850.

[36] X. Zhang, Q. Zhang, Monitoring interannual variation in global crop yield using long-term AVHRR and MODIS observations, ISPRS J. Photogramm. Remote Sens. 114 (2016) 191–205, https://doi.org/10.1016/j.isprsjprs.2016.02.010.

[37] P.L. Gregersen, A. Culetic, L. Boschian, K. Krupinska, Plant senescence and crop productivity, Plant Mol. Biol. 82 (6) (2013) 603–622, https://doi.org/10.1007/s11103-013-0013-8.

[38] D.B. Lobell, The use of satellite data for crop yield gap analysis, Field Crops. Res. 143 (2013) 56–64, https://doi.org/10.1016/j.fcr.2012.08.008.

[39] W. Masiza, G.J. Chirima, H. Hamandawana, A.M. Kalumba, H.B. Magagula, Do satellite data correlate with in-situ rainfall and smallholder crop yields? Implications for crop insurance, Sustainability 14 (3) (2022), https://doi.org/10.3390/su14031670.

[40] M.O. Román, C. Justice, I. Paynter, P.B. Boucher, S. Devadiga, A. Endsley, A. Erb, M. Friedl, H. Gao, L. Giglio, J.M. Gray, D. Hall, G. Hulley, J. Kimball, Y. Knyazikhin, A. Lyapustin, R.B. Myneni, P. Noojipady, J. Pu, R. Wolfe, Continuity between NASA MODIS Collection 6.1 and VIIRS Collection 2 land products. Remote Sensing of Environment, Elsevier Inc, 2024, https://doi.org/10.1016/j.rse.2023.113963, 302.

[41] K. Walker, Overcoming common pitfalls to improve the accuracy of crop residue burning measurement based on remote sensing data, Remote Sens. (Basel) 16 (2) (2024), https://doi.org/10.3390/rs16020342.

[42] T.E. Skosana, K.J. Esler, A.J. Rebelo, Exploring the trade-offs between spatial and spectral resolution in mapping invasive alien trees. Ecological Informatics, 2025, p. 92, https://doi.org/10.1016/j.ecoinf.2025.103448.

[43] M. Weiss, F. Jacob, G. Duveiller, Remote sensing for agricultural applications: a meta-review, Remote Sens. Env. (2020) 236, https://doi.org/10.1016/j.rse.2019.111402.

[44] M.A. Wulder, T.R. Loveland, D.P. Roy, C.J. Crawford, J.G. Masek, C.E. Woodcock, R.G. Allen, M.C. Anderson, A.S. Belward, W.B. Cohen, J. Dwyer, A. Erb, F. Gao, P. Griffiths, D. Helder, T. Hermosilla, J.D. Hipple, P. Hostert, M.J. Hughes, Z. Zhu,

Current status of Landsat program, science, and applications, Remote Sens. Env. 225 (2019) 127–147, https://doi.org/10.1016/j.rse.2019.02.015.
[45] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, P. Bargellini, Sentinel-2: ESA's optical high-resolution mission for GMES operational services, Remote Sens. Env. 120 (2012) 25–36, https://doi.org/10.1016/j.rse.2011.11.026.
[46] F. Gao, X. Zhang, Mapping crop phenology in near real-time using satellite remote sensing: challenges and opportunities, J. Remote Sens. (U. S.) (2021) 2021, https://doi.org/10.34133/2021/8379391.
[47] Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, 237. https://doi.org/10.1016/j.rse.2019.111511.
[48] D. Radočaj, I. Plaščak, M. Jurišić, Fusion of Sentinel-2 phenology metrics and saturation-resistant vegetation indices for improved correlation with maize yield maps, Agronomy (6) (2025) 15, https://doi.org/10.3390/agronomy15061329.
[49] D. Radočaj, I. Plaščak, M. Jurišić, Phenology-based maize and soybean yield potential prediction using machine learning and Sentinel-2 imagery time-series, Appl. Sci. (Switz.) (13) (2025) 15, https://doi.org/10.3390/app15137216.
[50] H.E. Beck, N.E. Zimmermann, T.R. McVicar, N. Vergopolan, A. Berg, E.F. Wood, Present and future köppen-geiger climate classification maps at 1-km resolution, Sci. Data 5 (2018) 1–12, https://doi.org/10.1038/sdata.2018.214.
[51] M. Nkamisa, S. Ndhleve, M.D.V. Nakin, A. Mngeni, H.M. Kabiti, Analysis of trends, recurrences, severity and frequency of droughts using standardised precipitation index: case of OR Tambo District Municipality, Eastern Cape, South Africa, Jamba: J. Disaster Risk Stud. 14 (1) (2022), https://doi.org/10.4102/jamba.v14i1.1147.
[52] C.K.E. Betek, N.D. Jumbam, Small scale vegetable production: a case study of Port St Johns, Eastern Cape Province, South Africa, Trans. R. Soc. S. Afr. 70 (1) (2015) 41–45, https://doi.org/10.1080/0035919X.2014.985759.
[53] Marais, M., & Swart, A. (2007). *Plant nematodes in South Africa. 8. Bizana, Lusikisiki and Port St Johns area, Eastern Cape Province*. https://doi.org/10.10520/EJC87805.
[54] J.L. Mwinga, W. Otang-Mbeng, B.P. Kubheka, A.O. Aremu, Ethnobotanical survey of plants used by subsistence farmers in mitigating cabbage and spinach diseases in OR Tambo Municipality, South Africa, Plants 11 (23) (2022), https://doi.org/10.3390/plants11233215.
[55] South African Government. (2013, April 3). *Swift actions taken to rectify the army worm outbreak*. https://www.gov.za/news/media-statements/swift-actions-taken-rectify-army-worm-outbreak-03-apr-2013.
[56] D. Blair, C.M. Shackleton, P.J. Mograbi, Cropland abandonment in South African smallholder communal lands: land cover change (1950-2010) and farmer perceptions of contributing factors, Land. (Basel) 7 (4) (2018), https://doi.org/10.3390/land7040121.
[57] K. Fischer, E. Johnson, V. Visser, S. Shackleton, Social drivers and differentiated effects of deagrarianisation: a longitudinal study of smallholder farming in South Africa's Eastern Cape province, J. Rural. Stud. 106 (2024), https://doi.org/10.1016/j.jrurstud.2024.103200.
[58] S. Pokwana, C. Shackleton, Complex and Interacting Drivers of Arable Field Abandonment in Villages of the Eastern Cape and KwaZulu-Natal, Land Use Policy, South Africa, 2025, p. 158, https://doi.org/10.1016/j.landusepol.2025.107736.
[59] COGTA. (2020). OR Tambo District Municipality: Profile and Analysis. https://www.cogta.gov.za/ddm/wp-content/uploads/2020/11/ORTamnco-September-2020.pdf.
[60] W. Masiza, J. Chirima, H. Hamandawana, R. Pillay, Enhanced mapping of a smallholder crop farming landscape through image fusion and model stacking, Int. J. Remote Sens. 41 (22) (2020) 8739–8756, https://doi.org/10.1080/01431161.2020.1783017.
[61] W. Masiza, J.G. Chirima, H. Hamandawana, A.M. Kalumba, H.B. Magagula, Linking agricultural index insurance with factors that influence maize yield in rain-fed smallholder farming systems, Sustainability 13 (9) (2021) 5176, https://doi.org/10.3390/su13095176.
[62] L. Mucina, M.C. Rutherford, The Vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute, South African National Biodiversity Institute, 2006. December 2015.
[63] M.E. Moeletsi, L. Myeni, L.C. Kaempffer, D. Vermaak, G. de Nysschen, C. Henningse, I. Nel, D. Rowswell, Climate dataset for South Africa by the agricultural research council, Data 7 (8) (2022) 1–12, https://doi.org/10.3390/data7080117.
[64] FAO. (2016). Crop Yield forecasting: Methodological and Institutional Aspects. https://doi.org/Available online: www.fao.org/publications (Accessed on 12 June 2025).
[65] W.S. Rezende, Y. Beyene, S. Mugo, E. Ndou, M. Gowda, J.P. Sserumaga, G. Asea, I. Ngolinda, M. Jumbo, S.O. Oikeh, M. Olsen, A. Borém, C.D. Cruz, B.M. Prasanna, Performance and yield stability of maize hybrids in stress-prone environments in eastern Africa, Crop J. 8 (1) (2020) 107–118, https://doi.org/10.1016/j.cj.2019.08.001.
[66] A. Tolera, F. Sundstùl, A.N. Said, The effect of stage of maturity on yield and quality of maize grain and stover, Anim. Feed Sci. Technol 75 (2) (1998) 157–168, https://doi.org/10.1016/S0377-8401(98)00192-8.
[67] GrainSA, Combat Harvest Time Losses, May 1, GrainSA, 2023, https://www.grainsa.co.za/combat-harvest-time-losses.
[68] AgBiz. (n.d.). *KNOWL. MODULE 4: Stored Grain Qual. Control NQF 5 8 Credits*. Retrieved January 20, 2026, from https://agbizgrain.co.za/content/openresource/9.
[69] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, Google Earth Engine: planetary-scale geospatial analysis for everyone, Remote Sens Env. 202 (2017) 18–27, https://doi.org/10.1016/j.rse.2017.06.031.
[70] J. Rouse, R. Haas, J. Schell, D. Deering, Monitoring vegetation systems in the Great Plains with ERTS, in: Proceedings of the Third ERTS-1 Symposium, NASA SP-351, 1974. https://ntrs.nasa.gov/api/citations/19740022614/downloads/19740022614.pdf.
[71] Z. Jiang, A.R. Huete, K. Didan, T. Miura, Development of a two-band enhanced vegetation index without a blue band, Remote Sens. Env. 112 (10) (2008) 3833–3845, https://doi.org/10.1016/j.rse.2008.06.006.
[72] A.A. Gitelson, M.N. Merzlyak, Remote sensing of chlorophyll concentration in higher plant leaves, Adv. Space Res. 22 (5) (1998) 689–692, https://doi.org/10.1016/S0273-1177(97)01133-2.
[73] C. Wu, Z. Niu, Q. Tang, W. Huang, Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation, Agric. Meteorol. 148 (8–9) (2008) 1230–1241, https://doi.org/10.1016/j.agrformet.2008.03.005.
[74] J.G.P.W. Clevers, A.A. Gitelson, Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and-3, Int. J. Appl. Earth Obs. Geoinf. 23 (1) (2013) 344–351, https://doi.org/10.1016/j.jag.2012.10.008.
[75] A.A. Gitelson, A. Viña, V. Ciganda, D.C. Rundquist, T.J. Arkebauer, Remote estimation of canopy chlorophyll content in crops, Geophys. Res. Lett. 32 (8) (2005) 1–4, https://doi.org/10.1029/2005GL022688.
[76] E.R. Hunt, B.N. Rock, Detection of changes in leaf water content using near- and middle-infrared reflectances, Remote Sens. Env. 30 (1) (1989) 43–54, https://doi.org/10.1016/0034-4257(89)90046-1.
[77] O. Kramer, Scikit-learn. Studies in Big Data, Springer Science and Business Media Deutschland GmbH, 2016, pp. 45–53, https://doi.org/10.1007/978-3-319-33383-0_5. Vol. 20.
[78] Haseeb, M., Tahir, Z., Mahmood, S.A., & Tariq, A. (2025). Winter wheat yield prediction using linear and nonlinear machine learning algorithms based on climatological and remote sensing data. *Information Processing in Agriculture*. https://doi.org/10.1016/j.inpa.2025.02.004.
[79] J. Ko, T. Shin, J. Ban, H.Y. Kim, Assessing maize growth and yield using a remote sensing–integrated crop model enhanced with lasso and ridge regression, J. Appl. Remote Sens. (02) (2025) 19, https://doi.org/10.1117/1.jrs.19.028506.
[80] L.E. Melkumova, S.Y. Shatskikh, Comparing Ridge and LASSO estimators for data analysis, Procedia Eng. 201 (2017) 746–755, https://doi.org/10.1016/j.proeng.2017.09.615.
[81] A. Altmann, L. Toloşi, O. Sander, T. Lengauer, Permutation importance: a corrected feature importance measure, Bioinformatics 26 (10) (2010) 1340–1347, https://doi.org/10.1093/bioinformatics/btq134.
[82] Lundberg, S., & Lee, S.-I. (2017). *Unified Approach Interpret. Model Predict.*. http://arxiv.org/abs/1705.07874.
[83] M. Dubova, S. Chandramouli, G. Gigerenzer, P. Grünwald, W. Holmes, T. Lombrozo, M. Marelli, S. Musslick, B. Nicenboim, L.N. Ross, R. Shiffrin, M. White, E.J. Wagenmakers, P.C. Bürkner, S.J. Sloman, Is Ockham's razor losing its edge? New perspectives on the principle of model parsimony, Proc. Natl. Acad. Sci. U.S.A 122 (5) (2025), https://doi.org/10.1073/pnas.2401230121.
[84] Di Bucchianico, A. (2008). Coefficient of determination (R2). *Encyclopedia of Statistics in Quality and Reliability*. https://doi.org/10.1002/9780470061572.eqr173.
[85] T.O. Hodson, Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geoscientific Model Development, Copernicus GmbH, 2022, pp. 5481–5487, https://doi.org/10.5194/gmd-15-5481-2022. Vol. 15, Number 14.
[86] C. Knight, A. Khouakhi, T.W. Waine, The impact of weather patterns on inter-annual crop yield variability, Sci. Total Environ. (2024) 955, https://doi.org/10.1016/j.scitotenv.2024.177181.
[87] D.K. Ray, J.S. Gerber, G.K. Macdonald, P.C. West, Climate variation explains a third of global crop yield variability, Nat. Commun. 6 (2015), https://doi.org/10.1038/ncomms6989.
[88] M. Kapari, L. Nhamo, M. Sibanda, J. Magidi, S. Mpandeli, T. Mabhaudhi, Remote sensing maize water stress in smallholder farms: a systematic review of progress, challenges, and the way forward using earth observation data. Enhancing Water and Food Security Through Improved Agricultural Water Productivity: New Knowledge, Innovations and Applications, Springer Science+Business Media, 2025, pp. 77–116, https://doi.org/10.1007/978-981-96-1848-4_4.
[89] Q. Liu, F. Yao, A. Garcia-Garcia, J. Zhang, J. Li, S. Ma, S. Li, J. Peng, The response and sensitivity of global vegetation to water stress: a comparison of different satellite-based NDVI products, Int. J. Appl. Earth Obs. Geoinf. 120 (2023), https://doi.org/10.1016/j.jag.2023.103341.