

High-resolution wall-to-wall time series of seasonal maize area and yield for Rwanda over 2019-2023

Katie Fankhauser^{1,2,4,*}, Evan Thomas¹, Christopher Brook³, Arsene Gatera³, and Zia Mehrabi^{1,2,4*}

¹Mortenson Center in Global Engineering & Resilience, University of Colorado Boulder, Boulder, CO, USA

²Better Planet Lab, University of Colorado Boulder, Boulder, CO, USA

³One Acre Fund, Kigali, Rwanda

⁴Department of Environmental Studies, University of Colorado Boulder, Boulder, CO, USA

*corresponding author(s): Katie Fankhauser (katie.fankhauser@colorado.edu) & Zia Mehrabi (zia.mehrabi@colorado.edu)

Abstract

Agricultural monitoring is least developed for smallholders in low- and middle-income countries — communities most likely to be impacted by hunger, poverty, and climate change. Recent efforts to monitor smallholder productivity are limited in spatial and temporal scope, but here, we provide an end-to-end machine learning pipeline built on Google Earth Engine for high-resolution, wall-to-wall time series mapping of crop area and yield, demonstrated for maize at every 10 m pixel in Rwanda over 2019-2023. Gradient boosted tree models were built from more than 60,000 field-level labels, 9,000 yield measurements, and satellite-derived inputs. Maize was classified with 83% accuracy, precision of 0.70, and recall of 0.44 and total maize cover was predicted within 4% of national statistics. Yields aggregated to districts had an RMSE of 370 kg/ha (27%). Our data compare favorably to other smallholder maize classification and yield estimation products for sub-Saharan Africa while being accessible, low-cost, standardized, and observed over time; thus, being more likely to enable technology transfer and innovation for many uses cases and users.

Keywords: agricultural monitoring, remote sensing, data and pipeline

1 Introduction

In 2022, around 740 million people, over 9% of the global population, faced a lack of basic calorie requirements, the majority of who reside in low- and middle-income countries (LMICs). Agricultural monitoring is key to understanding how food supply is changing and how stressed, agriculture-dependent communities are being affected by climate, economic, or conflict related shocks. Monitoring also provides a mechanism to evaluate how programs and policies aiming to improve agricultural outcomes are performing at scale [1]. Crop type and productivity data are critically important to agricultural monitoring systems, but are often non-existent, unstandardized, coarse, or proprietary [2]. Moreover, technology is least well developed for smallholders, the predominant agricultural system in LMICs. It is widely recognized by food system experts that there is a fundamental need to fill this gap and improve crop monitoring services in smallholder systems across the world [3].

Despite many recent advancements in satellite resolution and algorithmic development for crop monitoring, global models and systems will invariably exhibit substantial errors at the local scale, limiting their relevance for national and regional decision-making. This has prompted experts to call for custom on-demand maps for specific applications and regions [4].

One tool which offers promise is Google Earth Engine (GEE) and the access it affords to large geospatial data and distributed cloud-based computing. Methods built upon this framework have led to a proliferation of localized land cover and crop yield products. For example, GEE has been used in pipelines for predictions of crop type, and in some cases yield, for Nigeria [5], Ethiopia, Malawi [6], Rwanda [7], Tanzania, Kenya [8], and Zambia [4]. At the same time, the end-to-end pipelines for bringing these products closer to operational settings have been neglected and are often not accessible. While many authors discuss the benefit of scalable approaches that can be implemented across years and regions, the products they develop are nonetheless limited in scope to a single year or season and in most cases for sub-regions of a single or neighboring countries. Therefore, there is still considerable unmet potential to use GEE to develop high-resolution wall-to-wall seasonal time series predictions of crop productivity for specific contexts that are accessible and that lower the barrier for innovation for a variety of use cases and users.

Here, we document the development and implementation of such an end-to-end machine learning pipeline to predict maize cover and yield in Rwanda from 2019-2023 at 10 m resolution. We further illustrate how these products, with additional reference data, can be post-processed to conform with existing datasets and be consistent with national or inter-organizational statistics, which we have found can be essential for enabling decision-making at this level. We provide open access to the gridded datasets that represent crop type masks and yield outcomes as well as the code to produce these products. Our pipeline can be used to generate high-resolution estimates of seasonal maize production for an entire country within days and with nominal cost.

Rwanda provides an opportunity to demonstrate the development of a localized crop monitoring system and to test the utility of such technology and data. The government of Rwanda is committed to developing its agricultural sector [9], an industry that is critical to livelihoods and the national economy. Maize is a staple crop sensitive to short- and medium-term changes in the environment [10] that is of local importance to food security (contributing ~11% of food calories for the population) [11]. This setting offers the necessary conditions to better understand specific use cases, such as monitoring the response of local agriculture to climate or economic shocks, identifying the impact of new infrastructure on rural development, or the impact of strategies and policies promoted by governments and public, private, or civil society sector actors at scale.

In the following sections, we describe the methodology and validation of a data science pipeline that undertakes wall-to-wall land cover and crop type classification and yield mapping, with an emphasis on maize identification and yield for Rwanda, for two main growing seasons in each year 2019-2023 at 10 m resolution. We provide notes on access and use of the final data products and scripts while emphasizing the importance of developing accessible, localized time series for improving agricultural monitoring and innovation.

2 Methods

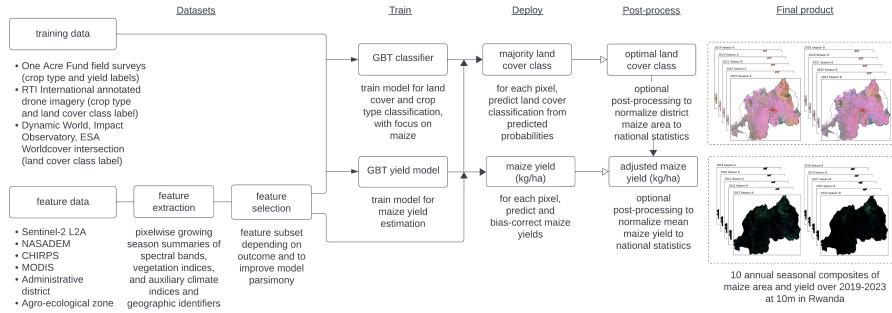


Figure 1: End-to-end machine learning pipeline for land cover classification with special emphasis on maize and maize yield prediction including data acquisition, pre-processing, model building (training), prediction (deployment), post-processing, and production of final products. GBT = gradient boosted tree

2.1 Study Area

The pipeline (Fig. 1) for local, dynamic maize classification and yield was demonstrated in Rwanda for 10 seasons over five years. Agriculture supports employment for 65% of the Rwandan population [12], mostly smallholder farmers on hillside and intercropped farms [12]. Maize is a staple crop in the country, produced on 16% of cultivated land [13] with a production value close to \$170 million USD annually [14]. Maize is typically planted and harvested during two distinct agricultural seasons per year: Season A from September to February and Season B from March to June [13]. There are 10 agro-ecological zones in Rwanda with distinct elevation, precipitation, and temperature regimes that exhibit important diversity in growing conditions [15]. Districts are the first level of decentralized government in Rwanda and are responsible for the execution of many socio-economic and agricultural policies [16], thus we also implemented several components of the pipeline at district level.

2.2 Datasets

In order to predict land cover and maize classification and maize yields across time and space, we relied on a range of datasets, including: field-level training data from several sources, optical imagery from satellite data, satellite-derived climate data, administrative boundary data, and national agricultural statistics.

All earth observation imagery, analysis, and outputs unless otherwise stated were hosted in the GEE environment [17] and developed with the Earth Engine Python API in Python v3.10.

2.2.1 Label data

Classification

Maize identification was the primary objective of our classification pipeline, but we also discriminated eight other land cover classes: non-maize annuals (legumes including bush and climbing beans, potatoes, sweet potatoes, and cassava), bananas as a non-maize perennial, scrub/shrubland, forest, wetland/flooded vegetation, water, built structures, and bare ground. Class labels were collated from several data sources, but field data was supplied by One Acre Fund (OAF), an organization supporting rural smallholder farmers in East Africa through improving access to farm inputs and markets, training, and financing.

OAF surveys farm productivity and management for a sample of farmers each year. Enumerators visited one of the farmer's fields and placed a harvest box at a randomly assigned number of steps along and into the field from the corner. GPS coordinates were recorded here. Observations were dropped if the GPS coordinates of the harvest box were duplicated due to enumerator error or if GPS accuracy exceeded 5 m. The harvest box was typically 5 m x 5 m when maize and/or beans were observed and 3 m x 3 m when the predominant crop was potatoes; however, other dimensions were less frequently used when the field was arranged in long, narrow terraces. The enumerator recorded the identity of the primary crop and whether intercropping was present. If the box contained more than one crop, the percent of primary crop and percent of intercropping were documented, but no other information was given for the secondary crop(s). Median self-reported field size was 0.05 ha. The OAF data provides approximately 36,000 crop labels from September 1st, 2020 to March 1st, 2023, but with most of the data being collected in the later seasons, 2022 Season B and 2023 Season A.

The OAF data was augmented by another approximately 2,200 class labels annotated from 3-4 cm drone imagery [18] collected by RTI International for Rwanda in 2019 Season A [19]. We harmonized the crop labels from this labeled drone imagery with the land cover classes described above.

Finally, to improve the representation of non-crop classes, which we found critical for discriminating land cover, we sampled the intersection of three existing global land use land cover (LULC) datasets so that the proportion of samples by class in our data was representative of the observed distribution of classes in the 2023 Season A Rwandan national agricultural statistics [13]. The three LULC layers that were used were Microsoft/Impact Observatory [20], Dynamic World [21], and WorldCover [22], each with a native spatial resolution of 10 m and summarized to the discrete class or mode for the year 2021. These are modeled products so their inclusion represents semi-supervision in non-crop land cover, but we attempt to overcome any uncertainty by only extracting training labels where all three agree in their classification.

Cumulatively, training data for the classification model included nearly 63,000 observations across the nine land cover classes and in 6 seasons (2019A, 2021A, 2021B, 2022A, 2022B, 2023A). The data was randomly split into training and testing sets, with 80% used for model parameter tuning and selection and 20% of the data completely held out for testing performance at the end of the pipeline.

Yield

The model for maize yield prediction was built from a subset of OAF training data where the predominant crop was maize and geolocated, field-measured yield was simultaneously collected. At

the harvest box, crop cuts were collected for maize and then bagged, dried, and the grain extracted before being weighed and reported as yield in kilograms per acre, which was later converted to kilograms per hectare (kg/ha). Training observations were censored to yields between 90 and 10,000 kg/ha (2% of observations in the initial data were below 90 kg/ha and a single observation was above 10,000 kg/ha, a likely data entry outlier). Setting the lower inclusion bound to 90 kg/ha avoids having to represent a mixed error distribution model, which would have been unstable given the few data points at zero, and defines crop failure as collection of less than one standard 90 kg bag [23], a severely low yield.

The training data for yield includes about 9,000 observations with the testing set comprised of 2,350 observations. The mean observed yield for maize in the training data was 3,680 kg/ha, which we find is greater than the latest national statistics and estimates by FAOSTAT - around 1,500 kg/ha [13, 24]. This discrepancy is well known in the country by different actors. For national decision-making it is important for predictions to align with expectations from other ministerial data. As such we address this in an optional post-processing step with a simple adjustment factor to harmonize the data with national surveys.

2.2.2 Features

Sentinel-2 data

The primary features we used for predicting maize classification and yield were derived from Sentinel-2 Level-2A atmospherically corrected surface reflectance, a product of the European Space Agency (ESA) Copernicus Program [25]. The twin satellites of the Sentinel-2 mission observe 13 bands in the visible, near-infrared (NIR), and short wave infrared (SWIR) optical spectrum at 10, 20, or 60 m resolution with a combined revisit frequency of 5 days. The Sentinel-1 satellite, also of the Copernicus Program, records Synthetic Aperture Radar (SAR) data at a similar spatial and temporal resolution. The use of radar data, which is impervious to clouds and haze, can be beneficial to land surface observation, but was not used in this study after preliminary analysis showed a deleterious effect on model performance. Similarly, [6] showed that SAR data did not contribute to significant gains in performance of their maize classifier in East Africa.

The Sentinel-2 cloud probability dataset [26], created with the s2cloudless algorithm, and near-infrared reflectance were used to mask clouds and cloud shadows from the images [27]. First, the Sentinel-2 L2A image collection [28] was filtered to images with cloud cover percentage less than or equal to 60% of pixels in the original metadata. Then, pixels representing clouds and their shadows were masked from the included images. Pixels with a cloud probability greater than 40% were considered clouds and cloud shadows were identified from the intersection of non-water dark pixels (i.e. an NIR value less than 0.15) and a hypothetical shadow projected at a distance of 1 km from cloud edges. Small clouds and shadows less than 2 pixels (20 m) were removed from the cloud layer and remaining cloud-shadow edges were morphologically dilated by 50 m. These final representations of clouds were used to create a time series of cloud-free Sentinel-2 images.

The independent candidate features used for modeling included raw values of each of the 13

multi-spectral bands, scaled by 10000, and a collection of derived indices. Several indices were intended to describe vegetation while others were designed to categorize bare soil, water content, and urban environments. See Table 1 for a list and explanation of each index.

Sentinel-2 images were downloaded from September 1, 2018 - the start date of the first agricultural season with available imagery from the L2A product over Rwanda - to July 1, 2023 - the end date of the last preceding agricultural season at time of development of our tool. Each image was cloud masked, scaled and used for the creation of indices, and then summarized to season by distribution and tertile at pixel level. The pixelwise minimum, maximum, mean, and standard deviation of each spectral band and index was taken over the respective season.

To capture phenology — the temporal signal of physiological stages of crop and vegetative development that can help resolve land cover types — the season was segmented into three equidistant, consecutive time periods and the median value of each band and index was taken in each [8]. The number of images available in a season to create the composites was also included as a feature to control for missing data due to cloud cover. Missing data was imputed with the seasonal mean from all other available images in the study period for that pixel. In total, 2228 cloud-free images were used to describe 10 agricultural seasons across five years at over 250 million pixels at 10 m resolution in Rwanda.

Auxiliary data

In addition to the time series observation of spectral images at the land surface, several static variables to improve classification and yield were considered. Including these in our final product was meant to categorize the environment-management-production interactions not captured by the signal in satellite-derived vegetation alone. Topographic features slope, elevation, and aspect at 30 m were downloaded from NASADEM, improved accuracy re-processing of Shuttle Radar Topography Mission data [29]. Administrative district [30] and agro-ecological zone [15] were translated to spatial data and pixels contained within the respective polygons were assigned their value. Predictions of yield distinctively considered growing season weather variables: seasonal means calculated from average daily land surface temperature at 1 km resolution from the MODIS satellite [31] and daily cumulative precipitation at 5 km resolution provided by CHIRPS [32].

Feature subset

Table 1 indicates which features were considered in classification and yield modeling. A total of 179 candidate features were available for prediction of land cover classification so we performed feature selection to improve model parsimony and efficiency. Similar to maize classification pipelines demonstrated in Azzari et al. [6] and Jin et al. [8], we removed features, ranked by their Mutual Information Score, that were highly correlated (i.e. greater than 0.8) with any other feature. We repeated this procedure until no two remaining variables had a correlation greater than 0.8 while retaining the variables with the highest Mutual Information Score in the set. A total of 69 features were selected with this approach and subsequently used in the classification model. For the prediction of yield, the 15 explanatory features included peak seasonal values of five select vegetation indices and descriptions of topographic, geographic, climatic, and confounding variables.

Table 1: Spatial and satellite-derived data used as explanatory features of land cover classification and maize yield

Feature	Name	Description	Source
B1	Aerosols	443 nm	Sentinel-2
B2	Blue	490 nm	Level-2A
B3	Green	560 nm	Sentinel-2
B4	Red	665 nm	Level-2A
B5	Red Edge 1	705 nm	Sentinel-2
B6	Red Edge 2	740 nm	Level-2A
B7	Red Edge 3	783 nm	Sentinel-2
B8	Near-Infrared Red (NIR)	842 nm	Level-2A
B8A	Red Edge 4	865 nm	Level-2A
B9	Water vapor	940 nm	Level-2A
B11	Short Wave Infrared (SWIR) 1	1610 nm	Sentinel-2
B12	Short Wave Infrared (SWIR) 2	2190 nm	Level-2A
NDVI*	Normalized Difference Vegetation Index	Most commonly used index to measure vegetative health, density, and canopy structure [6, 8]; $(B8 - B4)/(B8 + B4)$	Sentinel-2
kNDVI	kernel NDVI	Non-linear NDVI that exploits higher order relations between spectral bands and has better sensitivity to biophysical and physiological parameters [33]	Level-2A
NDRE1*	Normalized Difference Red Edge 1 Index	Red-edge exhibits greater stability and sensitivity to chlorophyll content and has been used to detect greening trend, chlorophyll concentrations in mid to late stages of growth, and crop productivity [34]; $(B8 - B5)/(B8 + B5)$	Sentinel-2
NDRE2	Normalized Difference Red Edge 2 Index	Another red-edge index that has similar advantages; $(B8 - B6)/(B8 + B6)$	Level-2A
GNDVI	Green Normalized Difference Vegetation Index	Sensitive to variations in chlorophyll content and photosynthetic activity that can be used to estimate maize yield [35] and assess distressed and aged vegetation [36]; $(B8 - B3)/(B8 + B3)$	Sentinel-2
GCVI*	Green Chlorophyll Vegetation Index	Sensitive to chlorophyll content, less likely to saturate at high leaf biomass, and correlated to leaf area index of cereal crops and crop yield [8, 35, 37]; $(B8/B3) - 1$	Level-2A
RV1	Ratio Vegetation Index	Sensitive to green vegetation and a good indicator of crop photosynthetic potential for sparse cereal canopies [38]; $(B8/B4)$	Sentinel-2
EVI*	Enhanced Vegetation Index	Remains sensitive to vegetative health and structure in dense canopy cover and less likely to saturate [39]; $2.5((B8 - B4)/(B8 + 6B4 - 7.5B2 + 1))$	Level-2A
REIP	Red-edge inflection point	Good indicator of crop photosynthetic potential for intermediate density cereal canopies and appropriate for field crop studies and monitoring [38, 40]; $702 + 35(((B4 + B7)/2) - B5)/(B6 - B5))$	Sentinel-2
NDTI	Normalized Difference Tillage Index	A "yellowness index" that monitors crop residues, canopy senescence, and grazing management [8, 41]; $(B11 - B12)/(B11 + B12)$	Level-2A
NDWI*	Normalized Difference Water Index	Sensitive to plant water content and demonstrated benefit for the remote sensing of maize yield [39]; $(B3 - B8)/(B3 + B8)$	Sentinel-2
NDBI	Normalized Difference Built-up Index	Separates urban from non-urban land cover by enhancing the spectral differences between SWIR and NIR observations [42, 43]; $(B11 - B8)/(B11 + B8)$	Level-2A
Precip_sum**	Precipitation	Cumulative daily precipitation	CHIRPS
LST_avg_mean**	Land surface temperature	Mean of daily daytime and nighttime land surface temperature	MODIS
District		Decentralized administrative level responsible for the execution of many socio-economic and agricultural policies	The World Bank
AEZ**	Agro-ecological zone	10 geographic zones with distinct elevation, rainfall, and temperature patterns	MINAGRI
Season_AB**		Binary indicator of whether maize yields were measured or predicted during a season A (Sept - Feb) or a season B (Mar - June)	MINAGRI
Num_imgs*		Number of Sentinel-2 imagery available in a season to create the composites	Sentinel-2
Longitude**		Longitude coordinate of pixel centroid	Level-2A
Latitude**		Latitude coordinate of pixel centroid	WGS84
Elevation*		Digital elevation model (DEM)	WGS84
Slope*		Slope, or steepness of the ground surface, in degrees calculated from the terrain DEM	NASADEM
Aspect*		Aspect, or compass direction that slope faces, in degrees calculated from the terrain DEM where 0=N, 90=E, 180=S, 270=W	NASADEM

Unless otherwise stated, all features were used to predict land cover classification with a special emphasis on maize identification: *Also used in prediction of maize yields, **Used exclusively for prediction of maize yields
 Sentinel-2 Level-2A [28], CHIRPS = Climate Hazards Group InfraRed Precipitation with Station [32], MODIS [31], The World Bank [30], MINAGRI = Rwanda Ministry of Agriculture and Animal Resources, NASADEM [29]

2.2.3 National agricultural statistics

As described above, under certain decision-making contexts, it is important to harmonize maize area and yield magnitude to national statistics. The National Institute of Statistics of Rwanda

(NISR) publishes agricultural surveys based on two-stage stratified population sampling around three months after the end of each main agricultural season [44]. Enumerators screen selected plots for land use, crop type, and area after planting and return after harvest to collect data on crop production. Enumerators delineate plot boundaries and visually inspect crop type at the point of data collection, but yield estimates are reported during interviews with farmers and are subject to respondent bias [13]. The seasonal surveys provide representative estimates of cultivated area and average yield by district for maize and other significant crops.

Thus, after performing prediction of maize cover and yield, we offer an optional product where predictions were normalized to the national statistics (Sections 2.3.1 and 2.4.1), which we recommend as the base product for most use cases. In addition, protected areas from the World Database on Protected Areas [45], such as preserved national parks, forests, and wetlands where no crop cultivation or monitoring is expected, were masked from the maize maps. These post-processing steps are described in more detail below. The final products provided here retain the raw predictions alongside the predictions optimized to be nationally representative to accommodate different downstream tasks.

2.3 Classification

A gradient boosted tree model, implemented with Tensorflow Decision Forests [46], was used to generate seasonally-explicit classifications of land cover, minimizing multinomial log likelihood loss in the nine-class outcome. The training set (which was 80% of the total initial labeled data; see Section 2.2.1) was further split using a 90:10 ratio and hyperparameter tuning (tree depth and shrinkage) was evaluated against the 10% set. The search space included 12 trials that represented possible tree depths up to 4, 5, 7, and 8 nodes and a shrinkage parameter of either 0.005, 0.01, or 0.02. Deeper trees increase complexity and contain more information about the interactions between predictors; Hastie et al. [47] suggest values between 4 and 8 for boosting methods. The shrinkage parameter is a regularization parameter applied to each tree prediction to control for overfitting; with a sufficient number of trees, smaller values are likely to give more accurate results. The number of iterations was set to 2,000, representative of 18,000 trees, to run the hyperparameter search. The best set of hyperparameters was informed by a balance between overall accuracy of the classifier, precision in the maize class, and computation time (e.g. of 12 candidate models, our final model had an accuracy and precision of 0.70 and 0.75 respectively, but required half of the computation time of the best model for accuracy and precision, at 0.72 and 0.74). A maximum depth of 5 nodes and a shrinkage rate of 0.01 were selected. We checked and confirmed that the models were not overfit with the validation dataset.

The final gradient boosted tree model was deployed on all training label data (i.e. no subsets reserved for validation) and used to predict land cover classification at all pixels in the region of interest in each season. The assigned class was taken as the class with the greatest predicted probability at each pixel centroid.

2.3.1 Optional post-processing

In our case, national agricultural statistics were available in each season (refer to Section 2.2.3) and so to align our predictions with these expectations for downstream tasks that would represent national outcomes, we created an optimal prediction for maize classification that used a varying class probability threshold to reduce error in predicted maize area within district and season between our product and national survey estimates.

We considered a set of possible probability class thresholds from 0 to 1 with a step size of 0.02. At each value, we calculated the total district area outside of protected lands that would be classified as maize given that the predicted probability was greater than the threshold. The threshold that resulted in the lowest absolute residual error in area compared to the official statistic was selected and deployed for that district and season. Since our pipeline can produce predictions immediately following the end of a season, in seasons where national data has not yet been published (usually several months post-season) we applied an optimal threshold value equal to the district mean in all previous A or B seasons respectively, which can easily be updated once the survey data becomes available. The optimally determined probability thresholds for maize classification ranged from 0.32 to 0.62 with a mean of 0.47.

2.4 Yield prediction

A similar modeling procedure described in Section 2.3 for land cover classification was followed for maize yield prediction. Again, a gradient boosted tree model and hyperparameter tuning of maximum tree depth and shrinkage rate was built from the subset of training data with yield information. Given that yield was a continuous measure, the model minimized squared error loss.

Hyperparameter tuning considered maximum tree depths between 4 and 8, shrinkage rates of 0.005, 0.01, and 0.02, and the optimal number of trees to avoid overfitting on a 10% validation set. The number of trees was initially set to 5,000 and training stopped for any given depth-shrinkage trial when the loss did not decrease in the next 30 look-ahead trees. The final chosen parameters were a maximum tree depth of 8 nodes, a shrinkage rate of 0.02, and 600 trees, which resulted in the lowest validation root mean squared error (RMSE) of all the trials. A gradient boosted tree model with these settings was fit to all training label data. We removed bias introduced by mis-estimation at the tails due to squared error loss by regressing predictions against the observations and forcing a 1:1 relationship between the two (Fig. A.1).

With the machine learning model we predicted yields in every pixel within Rwanda for each season over 2019-2023 and masked predictions to the maize class determined during classification (Section 2.3).

2.4.1 Optional post-processing

To align yield predictions to national statistics on maize yields (Section 2.2.3), we computed a simple adjustment factor equal to the ratio between the national mean predicted and mean reported yield in each season after masking to pixels predicted as maize in the post-processed optimal maize classification product. This season-specific adjustment factor, or the average historical adjustment

factor when official statistics are not yet available, was then applied to each pixelwise prediction. In the script included with the pipeline, there is an additional option to normalize mean yields by district instead of applying a single national factor, but we found that doing so could lead to artificial boundary effects between administrative units. The computed seasonal adjustment factors ranged from 2.2 to 2.9 with a mean of 2.5. This upward bias in yield was expected as mean yields in the training data were also around 2.5 times greater than national estimates (Section 2.2.1).

3 Results

3.1 Classification

For evaluating the land cover and maize classifications and maize yield estimation, we computed validation statistics on the 20% of data that was withheld from model building as a testing set. Table 2 contains validation statistics for both the training and testing sets, but since performance in the testing set was comparable to that in the training data and is a better measure of unbiased final prediction error, only results from the testing dataset are discussed below.

Across all nine land cover classes, the classifier had a balanced accuracy of 0.78. The majority class and optimal maize classifier similarly produced predictions with an F1 score of around 0.74 and an Matthew’s Correlation Coefficient (MCC) of at least 0.64, indicating favorable performance over the null with unbalanced data. For the singular maize class, accuracy was higher at 0.83 although the F1 score for this class was reduced to between 0.54-0.57, driven by poor recall as precision remained around 0.70.

High precision but low recall on the test labels is likely the result of intercropping and geospatially inaccurate labels (Table A.1). We find recall deteriorates when intercropping was present during the field survey, but remains low even in sub-field level monocropped observations, suggesting that the classifier has difficulty recognizing positive maize signals in intercropped sections of fields. Interestingly, when the performance of the model on OAF labels is compared to that on RTI labeled data created from high-resolution drone imagery, maize precision (0.85), recall (0.89), and accuracy (0.91) are much higher, indicating that there is most likely measurement error in the GPS points of most of the field data and we can achieve extremely high test performance on more precise labels. Moreover, the ability to identify maize in the noisier label data degrades with intercropping substantially. The classification of non-maize crops performs consistently well, but this was expected as the category represents a catchall of multiple other crops (beans, cassava, and potatoes) and does not require further discrimination.

Spatial inaccuracy in the OAF point data is reasonable as their primary data collection efforts inform program monitoring, evaluation, and learning and the data is being repurposed here for high resolution modeling. However, considering the uncertainty in the field training data lends more weight to the importance of scaling predictions to regional and national estimates. In fact, relative to national survey data, our majority class model over predicts maize area, which is the opposite of what might be expected under low maize recall from the model, indicating again the likelihood of noise in the labeled data from OAF. Normalizing predictions of maize cover to national statistics by

district did not impact classification performance compared to labeled data, but did greatly improve the RMSE from expected maize cover so that the optimal classification differs from total district maize cover by 4% or only \sim 200 ha on average.

3.2 Yield Prediction

The predictions of maize yields were different from observed yields on average by about 1650 kg/ha, or 44% of the mean in observed yields. The RMSE gives greater weight to outliers than the mean absolute error (MAE); when considering residual error irrespective of magnitude, average error in yield was 1300 kg/ha. The correlation between observed and predicted yield was moderate in the test set, with a value of 0.52 (and 0.65 in training).

In order to quantify the overall error expected when masking yield predictions to the upstream maize classification, we tested mean district yields from the national statistics against yield predictions at pixels identified as maize in the optimal class land cover classification. This considers the compound error from missed instances of maize as well as overestimating the presence of maize. Performance improved when maize yields were aggregated to district level. On average, the difference in mean district yield between the predictions and agricultural surveys was 374 kg/ha, or 27% of the overall mean, even when considering propagated error from the land cover classification product. A low mean error (ME) further supports aggregation and suggests that collectively when considering both under and overestimation of yield, we expect to exaggerate yield by only 50 kg/ha on average within our region.

4 Discussion

4.1 Comparisons to literature

Although sources of error exist in our maize classification and yield estimation products, performance is comparable to other efforts in the literature — maps that were produced at discrete time points and sub-regionally. Burke and Lobell [48] predicted maize cover in Kenya with 86% training accuracy and attained an R² coefficient of 0.20-0.39 for yield predictions using 1 m resolution imagery compared to our accuracy of 83% in the testing set (85% in training data) and R² value of 0.27 using lower resolution (10 m) Sentinel-2 data. In Ethiopia and Malawi, Azzari et al. [6] identified maize cultivation with 75% accuracy and found that some methods overestimated maize cover by 8-24% while our optimal maize land classification underestimates total maize area by only 0.06% on average per season. District level aggregation also improved performance of maize cover and yield prediction in Tanzania and Kenya [8]. The authors reported an accuracy between 63-79% for maize classification and developed a yield product with an R² of 0.54 and an RMSE of 670 kg/ha, but the normalized RMSE (nRMSE) relative to mean yields was roughly equal to 27% and comparable to our district-wise mean yield aggregation. The leading maize classification product in Rwanda is one produced by Hegarty-Craver et al. [7] which obtained an overall accuracy of 83% for land cover classification including maize, other crops, trees, other vegetation, and non-vegetation classes and a maize-specific precision of 85% and recall of 91% — this data was generated on a subset of that

used in our model and did not involve yield estimation. Our own training data includes these data [18] and our classifier can reproduce the high level of performance against these points (Table A.1), but they represent only a small fraction (3.5%) of the label data in our study with the majority of observations coming from ground-based surveying and crop cuts, which are subject to greater measurement error. In fact, many of the aforementioned studies incorporate training data that is produced by visual inspection and human annotation of high-resolution RGB images, which may account for their performance, relative to field surveys which may contain spatial inaccuracies in labels.

One of the key benefits of our product, which is comparable in performance to the best performing maize classification and yield products for sub-Saharan Africa and trained on a much larger training set, is that it provides a high-resolution time series of maize area and yield in all maize-growing agricultural seasons across several years and at every 10 m pixel for an entire country. Moreover, compared to traditional approaches, such as ground surveying and the generation of national agricultural statistics, our pipeline can produce these predictions within days after the end of a season. This approach should allow for the redistribution of time and resources away from the quantification of agricultural outcomes toward a focus on improving risk response and resilience in smallholder agriculture. Finally, we make not only the data product, but also all of the code readily available to allow others to replicate this work easily and undertake technology transfer to innovative use cases.

4.2 Final Product

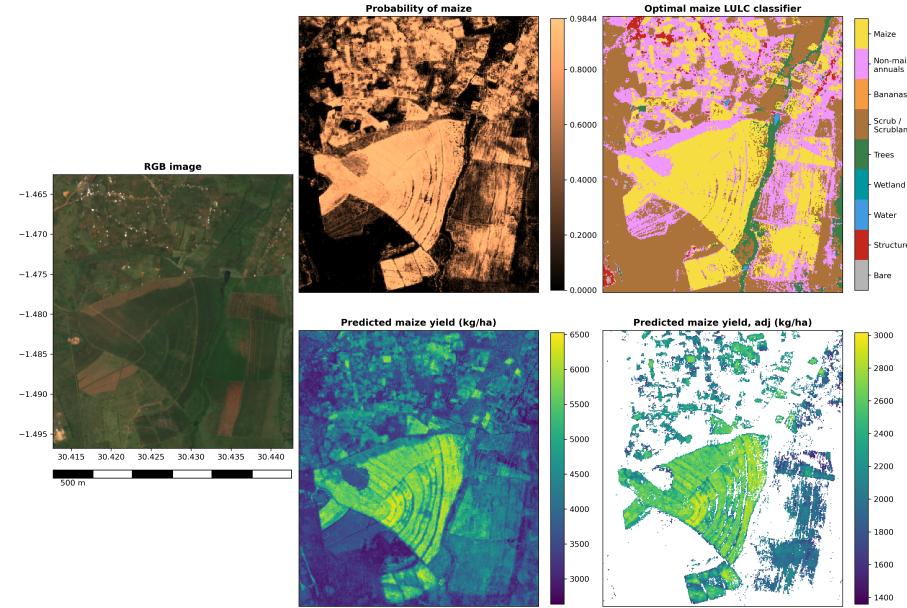


Figure 2: Demonstration of data products for maize classification and yield prediction with post-processing for a 3 x 3 km agricultural area in northeastern Rwanda.

The seasonal composites of 10 m resolution land cover and maize yield in Rwanda are hosted

Table 2: Validation statistics of classification and maize yield data products

	<i>Majority Class</i> train — test	<i>Optimal Class</i> train — test
Land Cover Classification		
Accuracy	0.836 — 0.786	0.834 — 0.784
F1 Score	0.787 — 0.744	0.779 — 0.734
MCC	0.709 — 0.650	0.700 — 0.641
Maize Classification		
Accuracy	0.856 — 0.832	0.852 — 0.827
F1 Score	0.624 — 0.570	0.600 — 0.541
Precision	0.755 — 0.691	0.760 — 0.692
Recall	0.531 — 0.485	0.495 — 0.444
District-wise*		
Maize cover area, RMSE (ha)	5285.677	199.881
Maize cover area, nRMSE	1.042	0.039

(a) Classification performance

	<i>Point Comparison</i> train — test	<i>District-wise Adj Yield*</i>
Maize Yield Estimation		
RMSE (kg/ha)	1414.525 — 1648.002	373.829
nRMSE	0.384 — 0.443	0.272
MAE (kg/ha)	1097.529 — 1300.537	292.035
ME (kg/ha)	0.525 — 51.186	-70.370
Pearson's Correlation	0.645 — 0.523	0.422

(b) Yield estimation performance

*District sums (maize cover) and means (maize yield) from national statistics compared against calculated district aggregations from final product layer, thus train and test sets are not applicable. Point comparison refers to labeled data from One Acre Fund against yield estimates regardless of land cover classification and unadjusted for scale. The district-wise adjusted yields represent the mean yield from optimal maize classification pixels that are scaled to national statistics.



Figure 3: Final data product of end-to-end machine learning pipeline for time series prediction and standardization of maize cover and yield in Rwanda. Shown are seasonal composites for each maize growing season over five years at 10 m resolution and detail for a 3 x 3 km agricultural area in northeastern Rwanda, the same area as in Fig. 2.

publicly on GEE and for download on Zenodo [49]. The repository contains detail on how to access these assets and includes the scripts and data used to create them. At the time of publication, each asset contained 10 images representing two annual agricultural seasons over five years (Fig. 3). To renew a season's worth of data would cost ~\$20 in GEE download and storage fees, a very low expense approach to timely agricultural monitoring for an entire country.

The land cover classification images contain 3 bands: *maizeProb*, the raw predicted probability of the pixel being maize given by the gradient boosted tree model; *majorityClass*, the categorical land cover class with the highest predicted probability among any of the nine classes in the respective pixel; and *optimalClass*, the categorical land cover class adjusted to agree with national statistics for expected maize area. Each of the images in the yield composites has 3 bands also: *maizeYield*, the model's output of continuous predicted yield (kg/ha) in each pixel regardless of land class; *maizeYield_majorityClass*, predicted maize yield masked to the majority class land classification; and *maizeYieldAdj_optimalClass*, where the raw predicted yields were masked to the optimal maize classification land cover layer and normalized to national statistics. Demonstration of maize probability, the optimal maize land cover classifier, and the raw and adjusted maize yields is given in Fig. 2 for a 3 x 3 km area in northeastern Rwanda. Ultimately, at the end of our pipeline, high spatial and temporal resolution standardized maize cover and yield (Fig. 3) are available for direct monitoring as well as secondary analyses.

We recommend using the optimized data products when studying regionally representative land cover and agricultural yield or comparing predictions to pre-existing aggregated data. Aligning predictions to the expectations of agricultural statistics produced by governments or multilateral institutions such as the FAO lends institutional validity and improves contextual relevance for partners working in the public sector. The data collected by these organizations expectedly differs from our training data (Section 2.2.3) and the variation has been quantified in post-processing. However, some measurement error is expected from both survey-based datasets and further insight into the actual magnitude of maize production would be gained by characterizing the source and distribution of error between the training data and national statistics.

The raw predictions were retained for users that wish to study relative change in maize cover or yield, the impact of policies implemented at district level (the unit of adjustment in our normalized predictions), or have other existing data to normalize to in their use case. Due to error in the unstandardized predictions, if users desire to use these, we do suggest aggregation, such as to the district level, for any population level inference. Since maize yield was predicted in every pixel regardless of land cover class, users of the raw predictions must mask yield to maize cover, using either the majority class product - a yield product that we have already generated and provided - or another maize cover layer provided by the user.

It is not recommended to use, or at least with extreme care, any of these products when studying the effect of any of the predictors directly. For instance, since precipitation and temperature were used to predict yield it may, depending on the application, be circular to use predictions to infer the downstream impact of climate on yield (first, the effects of temperature and precipitation in the yield outputs would need to be partialled out). The models should be applied outside of Rwanda with caution and only in places with similar environmental and agricultural systems. Instead, we

prefer that the pipeline be replicated in new contexts with relevant training data and retrained or fine-tuned models.

Our approach to ensuring data quality was to rely on ground data collected systematically by trained enumerators and standardized satellite-derived predictors and to institute data cleaning protocols, but there was likely some remaining noise in the training data. However, this work could not have proceeded without the large field dataset provided by OAF, making it critical to continue collecting ground truth data of this type as methods continue to develop despite the challenges [50]. Furthermore, the complexity of agricultural systems, including environmental drivers and human-induced factors such as socio-economics, farmer behavior, and cultural practices, make it difficult to achieve accurate and precise agricultural monitoring that explicitly incorporates these variables. The utility of the data presented here relies on adoption and uptake by stakeholders in scientific and implementing organizations. We have encouraged adoption by prioritizing ease of use, low-cost, and open access methods. Moreover, a proactive approach to continuous advancement of agricultural monitoring with new data and feedback improves not only the science but acceptance of new technological solutions.

This data enables a variety of exciting use cases. In an upcoming analysis, the authors will demonstrate the use of the optimal maize cover and yield predictions for localized monitoring of progress toward doubling agricultural productivity. It was these planned analyses that motivated the development and implementation of the maize cover and yield products, wherein we recognized their utility for other use cases and the importance of open access, reproducible science in agricultural monitoring.

5 Conclusions

We have presented an end-to-end machine learning pipeline and high-resolution time series for maize cover and yield in an agriculturally strategic country in East Africa. Our approach improves on global land cover and production products and demonstrates comparable performance to other regional products while being greater in scope both in terms of temporal and spatial resolution. Tailoring methods and datasets to address local challenges and maintaining a policy of open data and reproducibility exhibits a commitment to problem-solving with local stakeholders, practitioners, and other members of the scientific community. The value of this work will be enhanced if these methods and datasets lead to tangible improvements in agricultural monitoring and, subsequently, practices, decision-making, and policy.

Acknowledgments

Author Contributions

Conceptualization: K. Fankhauser, E. Thomas and Z. Mehrabi; Data curation: K. Fankhauser, C. Brook, and A. Gatera; Formal analysis: K. Fankhauser; Funding acquisition: E. Thomas; Investigation: K. Fankhauser; Methodology: K. Fankhauser, E. Thomas, and Z. Mehrabi; Project

administration: K. Fankhauser, E. Thomas, and Z. Mehrabi; Resources: E. Thomas; Software: K. Fankhauser; Supervision: E. Thomas and Z. Mehrabi; Validation: K. Fankhauser; Visualization: K. Fankhauser and Z. Mehrabi; Writing – original draft: K. Fankhauser and Z. Mehrabi; Writing - review & editing: K. Fankhauser, E. Thomas, C. Brook, A. Gatera, and Z. Mehrabi

Funding

This publication was made possible through support provided by the United States Agency for International Development under the terms of award number 7200AA20FA00021 and the Wellspring Foundation. The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of these institutions.

Conflicts of Interest

Co-authors C. Brook and A. Gatera are compensated employees of One Acre Fund, the organization that provided survey-based field data on crop cover and yield.

Data Availability

Scripts and data products described in this paper and additional usage notes can be found at the following repository <https://doi.org/10.5281/zenodo.10659095> [49].

Supplementary Materials

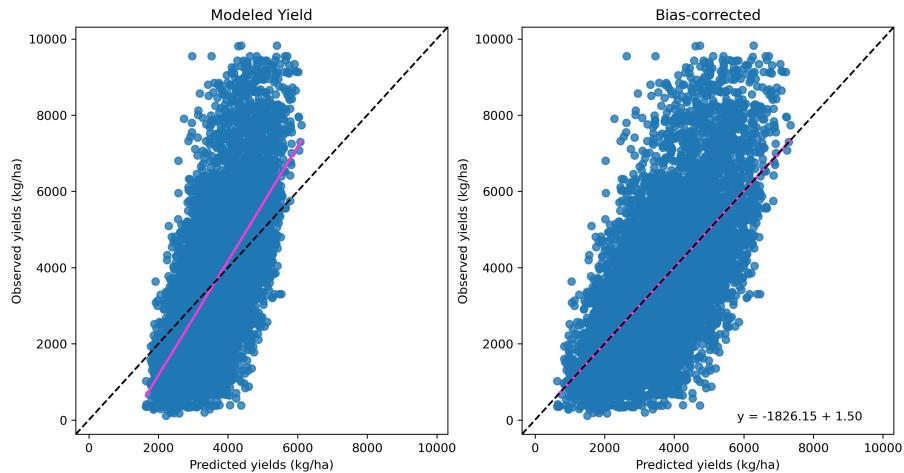


Figure A.1: Demonstration of bias correction in predicted yields

Table A.1: Classification performance in test set for crop cover by stand type and label data source

	Accuracy	F1 Score	Precision	Recall	N
Mono- vs Inter-cropped (OAF)					
<i>Monocropped</i>					
maize	0.737	0.587	0.795	0.465	1420
non-maize annual	0.732	0.793	0.737	0.859	2117
<i>Intercropped</i>					
maize	0.768	0.429	0.574	0.343	928
non-maize annual	0.741	0.831	0.808	0.855	2713
Dataset					
<i>OAF</i>					
maize	0.753	0.525	0.707	0.417	2348
non-maize annual (beans, potatoes, cassava)	0.737	0.814	0.775	0.857	4830
<i>RTI</i>					
maize	0.909	0.869	0.846	0.894	141
non-maize annual (legumes)	0.854	0.358	0.270	0.531	32
non-maize perennial (banana)	0.940	0.803	0.981	0.680	75

OAF = One Acre Fund field survey labels, RTI = RTI International annotated high-resolution drone imagery labels

References

1. Sachs J, Remans R, Smukler S, et al. Monitoring the world's agriculture. *Nature* 2010;466. Number: 7306 Publisher: Nature Publishing Group:558–60.
2. Fritz S, See L, Bayas JCL, et al. A comparison of global agricultural monitoring systems and current gaps. *Agricultural Systems* 2019;168:258–72.
3. Mehrabi Z, Delzeit R, Ignaciuk A, et al. Research priorities for global food security under extreme events. *One Earth* 2022;5:756–66.
4. Azzari G and Lobell DB. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sensing of Environment. Big Remotely Sensed Data: tools, applications and experiences* 2017;202:64–74.
5. Abubakar GA, Wang K, Koko AF, et al. Mapping Maize Cropland and Land Cover in Semi-Arid Region in Northern Nigeria Using Machine Learning and Google Earth Engine. *Remote Sensing* 2023;15:2835.
6. Azzari G, Jain S, Jeffries G, Kilic T, and Murray S. Understanding the Requirements for Surveys to Support Satellite-Based Crop Type Mapping: Evidence from Sub-Saharan Africa. *Remote Sensing* 2021;13:4749.
7. Hegarty-Craver M, Polly J, O'Neil M, et al. Remote Crop Mapping at Scale: Using Satellite Imagery and UAV-Acquired Data as Ground Truth. *Remote Sensing* 2020;12:1984.
8. Jin Z, Azzari G, You C, et al. Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sensing of Environment* 2019;228:115–28.
9. Cantore N. The crop intensification program in Rwanda: A sustainability analysis. Tech. rep. Overseas Development Institute, 2011. URL: https://rema.gov.rw/remadoc/pei/Report_ODI_AM.pdf (visited on 01/24/2024).
10. Chew A, Reynolds T, Waddington SR, et al. Maize Systems. Tech. rep. 215. Evans School Policy Analysis & Research Group (EPAR), 2013. URL: <https://epar.evans.uw.edu/research/agriculture-environment-maize-systems> (visited on 01/30/2024).
11. FAO. Food Balances (2010-). 2021. URL: <https://www.fao.org/faostat/en/#data/FBS> (visited on 01/04/2024).
12. MINAGRI. Ministry of Agriculture and Animal Resources Annual Report 2022/23. Tech. rep. Kigali, Rwanda: Republic of Rwanda, 2023. URL: <https://www.minagri.gov.rw/publications/reports> (visited on 01/08/2024).
13. NISR. Seasonal Agricultural Survey. 2023. URL: <https://www.statistics.gov.rw/datasource/seasonal-agricultural-survey> (visited on 01/04/2024).
14. FAO. Value of agricultural production. 2021. URL: <https://www.fao.org/faostat/en/#data/QV> (visited on 01/04/2024).
15. Nzeyimana I, Hartemink AE, and Geissen V. GIS-Based Multi-Criteria Analysis for Arabica Coffee Expansion in Rwanda. *PLOS ONE* 2014;9:e107449.

16. Republic of Rwanda. Government of Rwanda: Administrative Structure. 2024. URL: <https://www.gov.rw/government/administrative-structure> (visited on 01/09/2024).
17. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, and Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment. Big Remotely Sensed Data: tools, applications and experiences 2017;202:18–27.
18. Rineer J, Beach R, Lapidus D, et al. Drone Imagery Classification Training Dataset for Crop Types in Rwanda. 2021. DOI: 10.34911/rdnt.r4p1fr.
19. Chew R, Rineer J, Beach R, et al. Deep Neural Networks and Transfer Learning for Food Crop Identification in UAV Images. Drones 2020;4:7.
20. Impact Observatory. 10m Annual Land Use Land Cover (9-class). 2023. URL: <https://planetarycomputer.microsoft.com/dataset/io-lulc-9-class> (visited on 01/03/2024).
21. Brown CF, Brumby SP, Guzder-Williams B, et al. Dynamic World, Near real-time global 10 m land use land cover mapping. Scientific Data 2022;9:251.
22. Zanaga D, Van De Kerchove R, Daems D, et al. ESA WorldCover 10 m 2021 v200. 2022. DOI: 10.5281/zenodo.7254221. URL: <https://zenodo.org/records/7254221> (visited on 01/03/2024).
23. Dijkink B, Broeze J, and Vollebregt M. Hermetic Bags for the Storage of Maize: Perspectives on Economics, Food Security and Greenhouse Gas Emissions in Different Sub-Saharan African Countries. Frontiers in Sustainable Food Systems 2022;6.
24. FAO. Crops and livestock products. 2022. URL: <https://www.fao.org/faostat/en/#data/QCL> (visited on 01/04/2024).
25. ESA. Sentinel-2. 2023. URL: <https://sentiwiki.copernicus.eu/web/sentinel-2> (visited on 01/03/2024).
26. ESA. Sentinel-2: Cloud Probability. 2024. URL: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS%5C_S2%5C_CLOUD%5C_PROBABILITY (visited on 01/03/2024).
27. Braaten J. Sentinel-2 Cloud Masking with s2cloudless. 2020. URL: <https://developers.google.com/earth-engine/tutorials/community/sentinel-2-s2cloudless> (visited on 01/03/2024).
28. ESA. Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A. 2024. URL: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS%5C_S2%5C_SR%5C_HARMONIZED (visited on 01/03/2024).
29. NASA JPL. NASADEM Merged DEM Global 1 arc second V001. 2020. DOI: 10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001.
30. The World Bank. Rwanda Admin Boundaries And Villages. 2012. URL: <https://datacatalog.worldbank.org/search/dataset/0041453/Rwanda-Admin-Boundaries-and-Villages> (visited on 01/03/2024).

31. Wan Z, Hook S, and Hulley G. MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V061. 2021. doi: 10.5067/MODIS/MOD11A1.061.
32. Funk C, Peterson P, Landsfeld M, et al. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* 2015;2.
33. Camps-Valls G, Campos-Taberner M, Moreno-Martínez Á, et al. A unified vegetation index for quantifying the terrestrial biosphere. *Science Advances* 2021;7:eabc7447.
34. Misra G, Cawkwell F, and Wingler A. Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sensing* 2020;12:2760.
35. Zhang L, Zhang Z, Luo Y, Cao J, Xie R, and Li S. Integrating satellite-derived climatic and vegetation indices to predict smallholder maize yield using deep learning. *Agricultural and Forest Meteorology* 2021;311:108666.
36. Lima IP de, Jorge RG, and Lima JLMP de. Remote Sensing Monitoring of Rice Fields: Towards Assessing Water Saving Irrigation Management Practices. *Frontiers in Remote Sensing* 2021;2.
37. Ulfa F, Orton TG, Dang YP, and Menzies NW. Developing and Testing Remote-Sensing Indices to Represent within-Field Variation of Wheat Yields: Assessment of the Variation Explained by Simple Models. *Agronomy* 2022;12:384.
38. Broge N, Thomsen A, and Andersen P. Comparison of selected vegetation indices as indicators of crop status. *Geoinformation for European-wide Integration* 2003:591–6.
39. Bolton DK and Friedl MA. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology* 2013;173:74–84.
40. Salvoldi M, Tubul Y, Karnieli A, and Herrmann I. VENµS-Derived NDVI and REIP at Different View Azimuth Angles. *Remote Sensing* 2022;14:184.
41. Gan L, Cao X, Chen X, He Q, Cui X, and Zhao C. Mapping Shrub Coverage in Xilin Gol Grassland with Multi-Temporal Sentinel-2 Imagery. *Remote Sensing* 2022;14:3266.
42. Lynch P, Blesius L, and Hines E. Classification of Urban Area Using Multispectral Indices for Urban Planning. *Remote Sensing* 2020;12:2503.
43. Zha Y, Gao J, and Ni S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing* 2003;24:583–94.
44. NISR. Seasonal Agricultural Survey - 2023 Annual Report. Tech. rep. Kigali, Rwanda: The Republic of Rwanda, 2023. URL: <https://www.statistics.gov.rw/publication/seasonal-agricultural-survey-2023-annual-report> (visited on 01/04/2024).
45. UNEP-WCMC and IUCN. The World Database on Protected Areas (WDPA). 2023. URL: https://developers.google.com/earth-engine/datasets/catalog/WCMC_WDPA_current_polygons (visited on 01/04/2024).
46. Guillame-Bert M, Bruch S, Stotz R, and Pfeifer J. Yggdrasil Decision Forests: A Fast and Extensible Decision Forests Library. In: *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2023, Long Beach, CA, USA, August 6-10, 2023*. 2023:4068–77. doi: 10.1145/3580305.3599933.

47. Hastie T, Tibshirani R, and Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media, 2013.
48. Burke M and Lobell DB. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences* 2017;114:2189–94.
49. Fankhauser K, Thomas EA, Brook C, Gatera A, and Mehrabi Z. High-resolution wall-to-wall time series predictions of seasonal maize area and yield for Rwanda over 2019–2023. 2024. doi: [10.5281/zenodo.10659095](https://doi.org/10.5281/zenodo.10659095).
50. Burke M, Driscoll A, Lobell DB, and Ermon S. Using satellite imagery to understand and promote sustainable development. *Science* 2021;371:eabe8628.