Remote Sensing of Environment 228 (2019) 115–128

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Smallholder maize area and yield mapping at national scales with Google Earth Engine

Zhenong Jin[a,1], George Azzari[a,1,2], Calum You[a], Stefania Di Tommaso[a], Stephen Aston[b], Marshall Burke[a], David B. Lobell[a,*]

[a] Department of Earth System Science, Center on Food Security and the Environment, Stanford University, Stanford, CA 94305, United States of America
[b] One Acre Fund, Kigali, Rwanda

ARTICLE INFO

Keywords:
Sentinel-1
Sentinel-2
Data fusion
Cropland classification
Yield mapping

ABSTRACT

Accurate measurements of maize yields at field or subfield scales are useful for guiding agronomic practices and investments and policies for improving food security. Data on smallholder maize systems are currently sparse, but satellite remote sensing offers promise for accelerating learning about these systems. Here we document the use of Google Earth Engine (GEE) to build "wall-to-wall" 10 m resolution maps of (i) cropland presence, (ii) maize presence, and (iii) maize yields for the main 2017 maize season in Kenya and Tanzania. Mapping these outcomes at this scale is extremely challenging because of very heterogeneous landscapes, lack of cloud-free satellite imagery, and the low quantity of quality ground-based data in these regions.

First, we computed seasonal median composites of Sentinel-1 radar backscatter and Sentinel-2 optical reflectance measures for each pixel in the region, and used them to build both crop/non-crop and maize/non-maize Random Forest (RF) classifiers. Several thousand crop/non-crop labels were collected through an in-house GEE labeler, and thousands of crop type labels from the 2015–2017 growing seasons were obtained from various sources. Results show that the crop/non-crop classifier successfully identified cropland with over 85% out-of-sample accuracy in both countries, with Sentinel-1 being particularly useful for prediction. Among the cropped pixels, the maize/non-maize classier had an accuracy of 79% in Tanzania and 63% in Kenya.

To map maize yields, we build on past work using a scalable crop yield mapper (SCYM) that utilizes simulations from a crop model to train a regression that predicts yields from observations. Here we advance past approaches by (i) grouping simulations by Global Agro-Environmental Stratification (GAES) zones across the two countries, in order to account for landscape heterogeneity, (ii) utilizing gridded datasets on soil and sowing and harvest dates to setup model simulations in a scalable way; and (iii) utilizing all available satellite observations during the growing season in a parsimonious way by using harmonic regression fits implemented in GEE. SCYM estimates were able to capture about 50% of the variation in the yields at the district level in Western Kenya as measured by objective ground-based crop cuts.

Finally, we illustrated the utility of our yield maps with two case studies. First, we document the magnitude and interannual variability of spatial heterogeneity of yields in each district, and how it varies for different parts of the region. Second, we combine our estimates with recently released soil databases in the region to investigate the most important soil constraints in the region. Soil factors explain a high fraction (72%) of variation in predicted yields, with the predominant factor being soil nitrogen levels. Overall, this study illustrates the power of combining Sentinel-1 and Sentinel-2 imagery, the GEE platform, and advanced classification and yield mapping algorithms to advance understanding of smallholder agricultural systems.

1. Introduction

Smallholder farms are the most common form of agriculture in the world, particularly in food insecure regions of South Asia and Sub-Saharan Africa (Lowder et al., 2016; Samberg et al., 2016). For example, recent analysis of sub-national census data suggests that over

* Corresponding author.
E-mail address: dlobell@stanford.edu (D.B. Lobell).
[1] Joint lead authors.
[2] Current address: University of Edinburgh, School of Engineering, Edinburgh, EH9 3FB, UK.

https://doi.org/10.1016/j.rse.2019.04.016
Received 28 September 2018; Received in revised form 26 March 2019; Accepted 12 April 2019

90% of food calories in South Asia and 50% in Sub-Saharan Africa originate from farms smaller than 5 ha in size (Samberg et al., 2016). Yet much remains unknown about these smallholder systems, including basic information on how efficiently they use key inputs such as land and labor, and how these productivities compare to those on larger farms or the profitability of off-farm activities (Carletto et al., 2015; Mccullough, 2017). The need for improved data on smallholder systems has motivated renewed investment in rigorous ground surveys such as the World Bank's Living Standards Measurement Study – Integrated Surveys on Agriculture (Carletto et al., 2015), as well as interest in the potential use of satellite remote sensing to provide low-cost measures of some key aspects of smallholder farms (Burke and Lobell, 2017; Chivasa et al., 2017).

New satellite sensors with spatial resolutions of 10 m or finer have opened up new possibilities for monitoring smallholder fields. For example, a 1 ha field consists of only ~10 pixels in a Landsat image with 30 m resolution, making it extremely difficult to delineate fields. In contrast, that same field would contain ~100 pixels in an image from Sentinel-2 with 10 m resolution, or ~1000 pixels in an image from a Planetscope sensor with roughly 3 m resolution. Given the 5-day revisit of Sentinel-2 since early 2017 and the roughly daily re-visit of Planetscope since late 2017, it is now possible to observe smallholder fields at a much higher frequency than ever before.

However, several aspects of smallholder farming, beyond the small sizes of fields, are challenging in the context of remote sensing. First and foremost is the ubiquitous cloud cover in many smallholder regions, which reside predominantly in the tropics and are actively growing crops during the rainy season (which is also the cloudy season). Second is the tremendous heterogeneity of most smallholder systems, with families typically growing multiple crops, often interspersed with each other and often including tree crops such as avocado or mangoes. In addition to this within-farm heterogeneity, there are often large differences between the crops grown by smallholders in one region and by farmers in a neighboring district or state.

Despite these obstacles, recent work has shown that images from Sentinel-2, Planet, and other fine-resolution sensors can be used to provide yield estimates on individual fields for a range of crops including maize, wheat, and millet (Jain et al., 2016; Burke and Lobell, 2017; Jin et al., 2017a; Lambert et al., 2018). These examples were typically from small sub-regions of ~100 km^2 or less, often chosen based on the availability of reliable ground-based data on yields, such as from farmer surveys or on-farm crop cuts, as well as the existence of cloud-free optical data. Moreover, because these studies typically compared yield estimates for specific fields with ground-based data, they did not have to address one of the main challenges of heterogeneous landscapes, which is identifying where specific crops or mixture of crops are growing. An exception is Jin et al. (2017a), who performed a classification across 40,000 km^2 in the Western Kenya region using Sentinel-2 data in order to map maize areas and yields, which were then aggregated to the district scale for comparison with averages of on-farm crop cuts.

To make remote sensing a more reliable tool for measuring smallholder yields, a remaining challenge is to develop methods that are truly scalable to the national and global level. That is, they will work in any smallholder region in any year, with as little dependence as possible on obtaining cloud-free conditions or having access to reliable ground measurements for calibration. Here we describe the development of methods towards this end, building on prior work in Eastern Africa (Burke and Lobell, 2017; Jin et al., 2017a). The resulting approach is illustrated for the cases of Kenya and Tanzania, for which country-wide maps of maize yields at 10 m resolution are generated for 2017. We also illustrate two examples of how these types of datasets can help advance understanding of yield constraints in smallholder systems.

The amount of data required for this task is substantial, with the countries of Kenya (580,367 km^2) and Tanzania (947,303 km^2) together comprising over 1.5 million km^2, and with this study using both Sentinel-1 and Sentinel-2 data at 10 m resolution. Accounting for the 6-day revisit period of Sentinel-1 and 5-day revisit of Sentinel-2 at the equator, this translates to roughly 200 billion pixels of observations per month over the study region, or roughly 1 trillion over the course of a growing season. To access and process this volume of data, we rely on the unique capabilities of the Google Earth Engine (GEE) cloud-based platform (Gorelick et al., 2017).

The following section describes the data inputs used for this study, as well as a customized model for masking clouds and cloud shadows in Sentinel-2 for this region, given the relatively poor performance of the default cloud mask in this region. Section 3 describes the approach to mapping the presence of field crops in general, and maize in particular, which are important precursors to mapping maize yields in the region. Section 4 then presents an updated approach to mapping maize yields, based on combining crop model simulations with Sentinel-2 imagery, and compares estimates to available field data. Section 5 then applies the maize yields maps to two relevant questions about maize productivity in the region - how spatially variable are maize yields between farmers' fields, and which soil properties appear to be important factors in explaining this variation. Finally, Section 6 discusses the main lessons from this study and implications for future work.

2. Data inputs and preprocessing

2.1. Study area

Kenya and Tanzania are two of the largest maize producers in sub-Saharan Africa, ranking behind only Nigeria in terms of total maize area, with an official estimate of 2.1 and 4.1 Mha of harvested area and 1.5 and 1.4 t/ha of yield, respectively, in 2017 (FAO, 2018). The typical maize growing systems in both countries are smallholder fields, often intercropped with other staple or cash crops. Inputs level of hybrid seeds and fertilizers are surprisingly low, although intensification is happening for maize in particular (Sheahan and Barrett, 2017). Both countries span large climatic gradients and have a diversity of maize systems, including many regions with two rainy seasons, and hence two growing seasons per year. In this paper we focus on the main maize season for each country, which for Tanzania spans from roughly February to July, and in Kenya from March to September.

2.2. Imagery

East Africa is one of the many equatorial regions that are characterized by significant year-round cloud cover (Whitcraft et al., 2015). Clouds and other atmospheric phenomena can dramatically affect the quality of optical imagery. Although some techniques exist to identify and remove contaminated pixels (see below), relying exclusively on optical imagery in these regions may be challenging. For this reason, we employ both optical and radar data in this study. Synthetic Aperture Radar (SAR) sensors operate in a part of the microwave region of the electromagnetic spectrum that is insensitive to water vapor and thus unaffected by clouds or haze. The European Space Agency (ESA) recently launched a SAR and optical mission called Sentinel-1 and Sentinel-2, respectively. Both missions are part of a larger family of Sentinels that belongs to Copernicus, an Earth Observing program managed by the European Commission (see http://www.copernicus.eu/main/overview). All Sentinels are designed as two-satellite constellations to ensure a high revisit time and enough data coverage for a wide variety of applications. For all levels of processing of both sources of imagery, we relied completely on GEE.

The applications we present in this study all rely on phenological changes as a key feature. It is therefore important that the seasonality of each remotely-sensed pixel is captured in a consistent way across different locations and years, even though each pixel may contain observations from different dates. This problem is normally solved using

some compositing or reducing techniques. Here we applied two different approaches: 1) seasonal medians for classification problems, and 2) harmonic regressions for yield estimates, each of which are described below in Sections 3 and 4, respectively.

2.2.1. Sentinel-1

Sentinel-1 satellites carry a C-band SAR centered at the single frequency of 5.405 GHz and are capable of different data acquisition modes. We utilized exclusively the Interferometric Wide swath mode (IW), which is the standard mode over land (Torres et al., 2012). Although Sentinel-1 instruments are centered on a single frequency, IW images are acquired with dual polarization (VV and VH), resulting in two backscatter bands. The two-satellite configuration enables a 6-day revisit-time; however, this frequency of observation is currently available only in specific regions. Sentinel-1 data is pre-processed for standard noise corrections and to a resolution of 10 m before being ingested in the GEE data pool. Even with such corrections applied, Sentinel-1 imagery suffers from a form of noise called speckle-noise, which confers a random granular texture to each image. Speckle-noise is unique to SAR imagery and is caused by backscatter interferences between adjacent returns (Goodman, 1976). Although images affected by speckle-noise show large-scale distinguishable features, they may be seriously compromised at smaller scales (e.g. within an agricultural field in Eastern Kenya, which only comprise a few 10 m pixels). In fact, it has been shown how speckle-noise may create confusion in machine-learning algorithms (Dasari et al., 2015), a key component in this study. Among the several existing techniques to correct SAR imagery for speckle-noise, in this study we chose to apply an algorithm called the refined Lee filter, as described in Abramov et al. (2017).

2.2.2. Sentinel-2

Sentinel-2 satellites carry an optical sensor called the Multi-Spectral Instrument (MSI) which acquires data from 13 spectral bands of variable resolution (see Table 1). In addition to the standard bands, we computed a set of vegetation indices (VIs) that have been used in the literature for a variety of applications. A summary of the VIs and their definitions are reported in Table 1. The two-satellite constellation enables a 5-day revisit-time. ESA distributes online only one official product for Sentinel-2, called Level-1C and consisting of Top Of Atmosphere (TOA) reflectance observations in cartographic geometry.

A second, higher-level product, called Level-2A consists of Surface Reflectance (SR) in cartographic geometry, and can be generated from Level-1C imagery using a dedicated Toolbox provided by ESA (https://github.com/senbox-org). Currently, however, Level-2A imagery is not available for download as a pre-generated product from either ESA or GEE. Because ESA's Toolbox is not directly available in GEE, the only option to use Level-2A in GEE is to pre-compute the product and ingest it in GEE. Such an approach, both computationally and storage-wise demanding, does not scale well to studies of large spatial extents such as presented here. Although not essential for the crop classification part of this study, where relative spectral differences are the essential aspect, SR imagery remains important for crop yield estimates. We therefore adopted a simple empirical approach for relating Level-1C imagery available in GEE to SR in the region, as explained in more detail in Section 4.2.

2.2.3. Cloud mask

Although ESA delivers Sentinel-2 Level-1C imagery with a Quality Assessment (QA) band that flags clouds and cirrus pixels, we observed misclassifications in several images over our study area. Furthermore, the native QA band does not provide information on haze and shadows. ESA's Toolbox provides a more advanced cloud detection algorithm. However, as noted above the scale of our study makes it difficult to run the Toolbox for the whole collection of imagery used. Other studies have shown promising results using a decision tree or Random Forest classifier to detect contaminated pixels (Hollstein et al., 2016). A key characteristic of decision-tree classifiers is that they can be completely serialized in string format and re-used to make predictions in other contexts, without needing to re-train the model. Such portability is a major advantage in this context, as it allows us to generate scalable predictions through GEE. Hollstein et al. (2016) provided decision trees for cloud, cirrus, shadows, and snow. A preliminary analysis in our area of interest showed some misclassification in masks generated using the Hollstein tree (HLS), particularly for the cloud and cirrus classes. This finding does not necessarily reflect poorly on Hollstein et al. (2016), since their model was trained globally and performance of a global model in any particular region will often exhibit some errors.

Nonetheless, in order to obtain a model tailored for our area, we collected training samples manually, by labeling Sentinel-2 images where clouds, haze, or shadows occurred. We selected training images by visual inspection and based on their location and time of the year. We labeled a total of 3,568 polygons over Sentinel-2 images from 65 different dates in 2016 and 2017, using 70% of these (2497) for model training and validation and 30% (1071) for testing and evaluation. For model selection, we sampled 15 points within each polygon which resulted in a total of 53,498 point-image observations after removing those with missing band values. We trained the decision tree offline, using the R package "rpart" (Therneau and Atkinson, 1997). The model performance is summarized in Fig. S1, with a comparison to both the QA band of Sentinel-2 and the HLS model. Overall, the model performed well with a test accuracy of 80% and a kappa score of 0.74, as compared to an accuracy of 63% for HLS and 55% for the QA band. Our model had some tendency to over-classify shadows (often confused with clear dark targets, such as burn scars or dark vegetation), while the HLS was more accurate in shadow detection. Given the overall better performance of our mask, it was used in all subsequent analyses to remove contaminated pixels in Sentinel-2 imagery.

2.3. Field data

Field data for the maize classification and yield validation tasks were obtained from several sources. These data had a spatial information component, which could either be boundary polygons of fields or simple points taken from within fields. Data providers linked these spatial geometries to attributes, which would be an observation of the crop type in a field and any intercropping if present, and could additionally include yields measured with a series of crop cuts. The main data providers are summarized below and in Table 2:

- One Acre Fund (1AF), a non-profit providing smallholder farmers with guidance and inputs, was the primary source of geolocated maize yield data, collected in both 2016 and 2017 in the Western province of Kenya. Crop cuts were conducted for ~4,000 fields in each year, with a GPS point collected in the middle of each field. On arriving at each field to be sampled, the enumerator randomly generated a number from 1 to 4 using their tablet to determine which corner of the field to go to. The enumerator then walked 4 steps along the row and 4 steps into the field to set the starting point for the first crop cut. The second crop cut was placed 8 steps towards the opposite corner, from the corner of the fist plot which is closest the center of the field. Placement of the plot may be adjusted slightly to avoid any obstacles such as termite mound and trees in the fields, although this happened very occasionally during the campaign. The size of crop cut plot is $5 \times 8\ m^2$. Harvested maize was stored in labeled bags for each crop cut, then dried, shelled, threshed and winnowed before enumerators returned for final weight. Although crop cuts are intended to be representative for each field, the low GPS accuracy (usually off by 10–30 m) on the ground prevents us from comparing crop cuts with satellite data specific to individual fields. For this reason, the crop cut data were aggregated up to the district level as in Jin et al. (2017a) and compared to satellite estimates at that scale.

Table 1
Bands and spectral indices from Sentinel-1 and Sentinel-2 data.

Band or index	Central wavelength/index formula	Satellite
VV	Vertically polarized backscatter	Sentinel-1
VH	Horizontally polarized backscatter	Sentinel-1
VV_RLSPCK	VV with refined Lee speckle filter	Sentinel-1
VH_RLSPCK	VH with refined Lee speckle filter	Sentinel-1
RATIO	VH/VV	Sentinel-1
DIFF	$VV - VH$	Sentinel-1
RATIO_RLSPCK	VH_{RLSPCK}/VV_{RLSPCK}	Sentinel-1
DIFF_RLSPCK	$VV_{RLSPCK} - VH_{RLSPCK}$	Sentinel-1
AEROS	443 nm	Sentinel-2
BLUE	490 nm	Sentinel-2
GREEN	560 nm	Sentinel-2
RED	665 nm	Sentinel-2
RDED1	705 nm	Sentinel-2
RDED2	740 nm	Sentinel-2
RDED3	783 nm	Sentinel-2
NIR	842 nm	Sentinel-2
RDED4	865 nm	Sentinel-2
VAPOR	940 nm	Sentinel-2
CIRRU	1375 nm	Sentinel-2
SWIR1	1610 nm	Sentinel-2
SWIR2	2190 nm	Sentinel-2
NDVI	$(NIR - RED)/(NIR + RED)$	Sentinel-2
RDNDVI1	$(NIR - RDED1)/(NIR + RDED1)$	Sentinel-2
RDNDVI2	$(NIR - RDED2)/(NIR + RDED2)$	Sentinel-2
GCVI	$(NIR/GREEN) - 1$	Sentinel-2
RDGCVI1	$(NIR/RDED1) - 1$	Sentinel-2
RDGCVI2	$(NIR/RDED2) - 1$	Sentinel-2
MTCI	$(NIR - RDED1)/(RDED1 - RED)$	Sentinel-2
MTCI2	$(RDED2 - RDED1)/(RDED1 - RED)$	Sentinel-2
REIP	$700 + 40 * \left(\frac{RED + RDED3}{2} - RDED1\right)/(RDED3 - RDED1)$	Sentinel-2
NBR1	$(NIR - SWIR1)/(NIR + SWIR1)$	Sentinel-2
NBR2	$(NIR - SWIR2)/(NIR + SWIR2)$	Sentinel-2
NDTI	$(SWIR1 - SWIR2)/(SWIR1 + SWIR2)$	Sentinel-2
CRC	$(SWIR1 - GREEN)/(SWIR1 + GREEN)$	Sentinel-2
STI	$SWIR1/SWIR2$	Sentinel-2

Table 2
Description of ground data sources, including the country, provider and year that they pertain to, whether point data or field polygons were supplied, the number of resultant points after area-based sampling, and whether the dataset was used for training in the crop type (maize/non-maize) classification task and/or validation of yield predictions.

Country	Provider	Year	Geometry	Points	Crop type task	Yield task
Kenya	1AF	2016	Points	3944		Yes
Kenya	1AF	2017	Points	4202		Yes
Kenya	IPA	2017	Polygons	16,964	Yes	
Tanzania	AFSIS	2015	Points	2075	Yes	
Tanzania	AFSIS	2016	Points	7999	Yes	
Tanzania	AFSIS	2017	Points	35	Yes	
Tanzania	TAMASA	2015	Points	136	Yes	
Tanzania	TAMASA	2016	Polygons	2383	Yes	
Tanzania	TAMASA	2017	Polygons	3219	Yes	

- Taking Maize Agronomy to Scale in Africa (TAMASA), a research consortium focused on raising productivity and profitability for maize farmers in several African countries, provided field polygons for both maize and non-maize fields throughout Tanzania.
- The Africa Soil Information Service (AFSIS) collected presence/absence data on 7 land cover types in a recent soil survey across Tanzania. These data are publicly available at https://github.com/mgwalsh.
- Finally, we contracted Innovations for Poverty Action (IPA) to conduct a survey in Kenya in 2017, to delineate field polygons for various crop types in the Western province.

In processing this data, we first collapsed crop type data into binary classes for maize classification. We treated intercropped maize as the maize class since maize is usually grown with lower-height crops that have a smaller impact on the spectral signature. To generate training points for maize classification, we combined point data with points sampled from within field polygons. The number of sampled points depended on the field area; small fields $< 800\,m^2$ in size had only their centroid sampled, larger fields with $> 2000\,m^2$ had five points randomly sampled, and those in between had 2–4 points sampled depending on the area. This was intended to limit duplication of sampled pixels in the dataset as well as overrepresentation of particularly large fields. Finally, although all classification for this study was done using 2017 imagery, we included field data on crop locations from 2015 and 2016 because the location of crops is generally stable from year to year, and the out-of-sample classification accuracies were generally lower when including more training data, even if from a different year. The dataset properties, number of points and the tasks they were used for are summarized in Table 2 and shown in Fig. 1.

In addition to the ground data on crop locations and yields, we also required data on locations of croplands vs. other land cover types. Leveraging GEE's User Interface (UI) API, we built an interactive platform to label random grids of points within a prescribed area based on the Code Editor basemap (which is the same as the very-high-resolution satellite imagery seen in Google Maps). The authors labeled these points based on visually interpreting the basemap.

Specifically, the labeling platform lets users select a country or a subregion of interest from a list, in this case, we included all admin Level-1 boundaries for Kenya and Tanzania, as downloaded from the GADM website (https://gadm.org/). Once the region is selected, the platform generates a prescribed number of random points within that region and iteratively visualizes each point on the basemap at a high zoom level, prompting users to select one of five classes available or

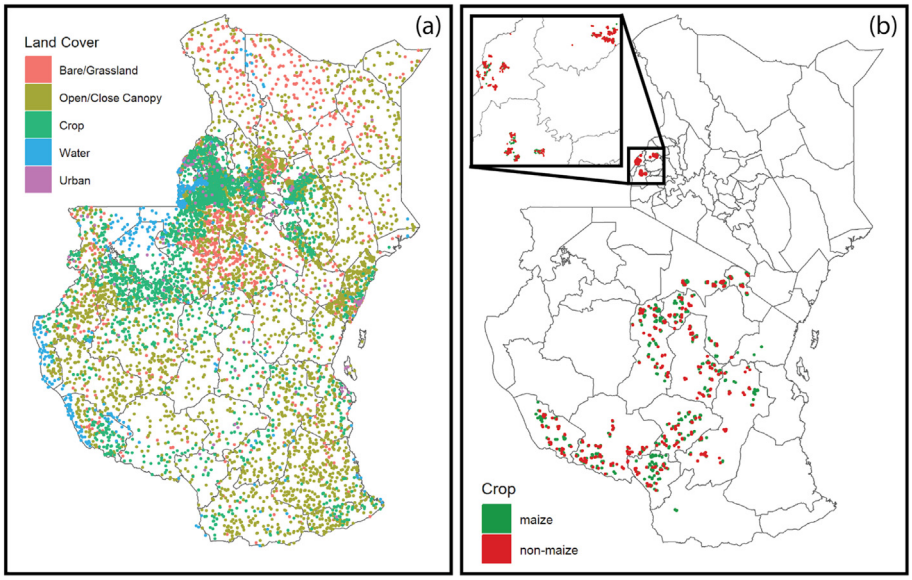

Fig. 1. Study area of Tanzania and Kenya showing (a) locations of training data for crop/non-crop classification, (b) location of training data for maize/non-maize classification.

skip to the next point. We defined the classes based on five main groups of cover types we observed in the area, namely 1) bare soil and grassland, 2) savanna, 3) crop, 4) water and swamps, 5) urban. These classes were determined by visual assessment and based on our previous experience in sub-Saharan Africa (Azzari and Lobell, 2017; Jin et al., 2017a). Once the user has labeled all points, the resulting geolocated grid with the corresponding labels is exported in KML format. We collected in total 4,140 labels in Tanzania and 4,509 labels in Kenya, the distribution of which is shown in Fig. 1.

3. Crop classification

We approached the goal of identifying maize pixels as split into two tasks: one to distinguish cropland pixels (CRL) from non-crop pixels, and one to identify maize pixels (MZL) from other types of crop pixels. Although the final goal was to generate the MZL map, we chose to first generate a CRL map to isolate crop pixels and reduce the overall intra-class variance in the MZL classification process.

3.1. Seasonal median composites

Phenological features have been proven to be a key feature in resolving different cover types. However, including phenology in a model requires metrics that measure seasonal changes consistently from year to year and place to place. In this study, we generated seasonal median composites (Figs. S2 and S3), an approach that has been successfully used in other studies spanning a variety of classification problems (e.g. Azzari and Lobell, 2017). In this approach, images from both Sentinel-1 and Sentinel-2 collections were divided into three groups based on their acquisition dates. The pixel-level median of each group is taken, each band of the resulting composite is renamed by adding a season-related suffix (e.g. “NDVI” of seasonal group 1 becomes “NDVI-S1”), and the all bands from the three composites are stacked together into a single image. Assuming that each image in the source collection had N bands, the resulting seasonal composite will have $N \times 3$ bands; this translated to a total of 24 features for Sentinel-1, and 69 features for Sentinel-2. Rather than using a traditional calendar year, we used a “water year” calendar spanning from October to September and grouped the imagery for as follows: 1) from October 2016 to January 2017, 2) from February 2017 to May 2017, and 3) from June 2017 to September 2017.

3.2. Classification pipeline

We used the same architecture for both CRL and MZL classification pipelines, which consisted of two main steps, 1) pre-selection of features, and 2) classification.

3.2.1. Feature pre-selection

The seasonal compositing approach we adopted as model inputs yielded a large number of features. Having a high number of features in a classification problem can lead to overfitting, become harder to interpret, and require larger computational resources. Furthermore, input features are often highly correlated with each other (e.g. a VI and its individual components), increasing the model's complexity while only providing redundant information. To subset the number of features as input to the classification model, we adopted a data-driven approach consisting of two phases. In the first phase, we computed the Mutual Information score (MI) for all features against the labels. MI is a metric that measures relative dependence between two variables even in non-linear cases and between continuous and categorical variables (Kraskov et al., 2004; Ross, 2014). In the second phase, we computed a correlation matrix for all features sorted by MI score, and iteratively dropped any feature that had a correlation higher than 0.8 with another higher-ranking feature. The remaining features were then passed to the model cross-validation pipeline described below. A similar methodology has been used by Chang et al. (2007) to selected features for mapping corn and soybean in the US. To assess the importance of each Sentinel satellite, we repeated the selection in three ways: 1) Sentinel-1 alone, 2) Sentinel-2 alone, and 3) Sentinel-1 and Sentinel-2 combined.

3.2.2. Model selection

We chose a Random Forest (RF) classifier for both CRL and MZL maps. RF classifiers have been successfully used in a variety of satellite-based applications including crop type mapping, and have the valuable advantage of being intrinsically not prone to overfitting (Inglada et al., 2015; Belgiu and Drăguţ, 2016). Importantly, an implementation of RF is available in GEE, allowing large-scale predictions at the pixel level. As with many other models, RF classifiers are sensitive to the choice of hyperparameters and the training data used. To construct the best possible CRL and MZL models, we utilized a K-fold cross-validated grid search. In this approach, the dataset is first randomly split into a validation (80% of the total dataset) and a test set (20% of the total

dataset). While the test set is kept aside for a final evaluation, the validation set is further split into K subsets (or folds). Splitting into folds can be done following specific criteria, such as using some categorical variable in the dataset (e.g. the district name) or completely randomly. Given a prescribed set of hyperparameters, a model is iteratively trained on K-1 of the folds and validated on the remaining one. The operation is repeated for any combination of hyperparameter values spanning a prescribed range of values. A set of performance metrics for both training and testing is recorded at each iteration and at the end of the process a grid with the full set of scores and the corresponding hyperparameters are exported. The model with the minimum number of features among the best-performing ones is used to run predictions in GEE. We repeated the model selection pipeline independently for each country and, to assess the importance of each Sentinel satellite, we repeated the selection in three ways: 1) Sentinel-1 alone, 2) Sentinel-2 alone, and 3) Sentinel-1 and Sentinel-2 combined.

For both CRL and MZL, we varied 4 hyperparameters during cross-validation: 1) n_estimators - the number of trees in the RF classifier, 2) max_features - the method used to determine the number of features considered at each tree split, 3) min_samples_split - the minimum number of samples required to split each node, and 4) min_samples_leaf - the minimum number of samples required to be at a terminal leaf node. We used two different types of folding criteria for CRL, 1) a random split across the whole validation dataset, and 2) a district-based split, to ensure that test folds were geographically distinct from the training folds. For MZL, given the clustered spatial distribution of the training data, we only used a random split by field. For both classification tasks, we used the hyperparameters n_estimators = 300, max_features = 'sqrt' and min_samples_leaf = 10 when implementing the model in GEE.

3.3. Classification results

3.3.1. Cropland classification

The subset of features selected for the CRL task for both Kenya and Tanzania are presented in Table S1, and the ranked correlation matrix before and after the iterative pruning process is shown in Fig. S2. The features selected for Kenya and Tanzania were different, with the only common feature among the 10 highest-ranking ones was Longitude. Four of the top-10 features in Kenya were "greenness indices" such as NDVI-S1, RDGCVI1-S2, and RDGCVI2-S3. In Tanzania's case, shortwave infrared dominated among the first ten features, with SWIR1-S2, SWIR1-S3, and SWIR2-S1 ranking first, second, and fifth, respectively. Furthermore, STI-S1 and NDTI-S3, two indices based on SWIR bands, ranked sixth and ninth. Sentinel-1 bands also played an important role, with two bands within the top-ten features of both countries. In Kenya, the speckle-corrected version of the VH/VV ratio for Season-2 and Season-3 (RATIO-RLSPCK-S2, and RATIO-RLSPCK-S3) ranked sixth and seventh, respectively, while in Tanzania RATIO-RLSPCK-S1 and VH-RLSPCK-S1 ranked 7th and 4th. The selection of different features for the two countries could arise for many reasons, but two likely ones are: (i) differences in the availability of cloud-free optical observations, with TZA relying more on Sentinel-1 SAR data because of fewer cloud-free optical images, and (ii) more distinct differences in soil wetness between crop/non-crop landscapes in Tanzania, which also increases the value of Sentinel-1.

For both countries, the accuracy of the model using both Sentinel-1 and Sentinel-2 was roughly 85% for the test dataset (i.e. the 20% of points that were not used in training). The accuracy of the model appeared to increase with the number of features until roughly the 20th feature in each country (Fig. 2). The hyperparameters of the best model were used to set up and run an RF classifier in GEE; the resulting cropland mask is shown in Fig. 3.

3.3.2. Maize classification

Table S2 presents the features selected for the MZL classification, along with their MI scores. Fig. S5 displays the ranked correlation matrix before and after the feature selection process. In addition to the 0.8 correlation threshold used to discard features, we also removed LON and LAT from the final model due to visual artifacts created in the final classification map. Several features were important for both the Kenya and Tanzania models, including AEROS_S1, VAPOR_S2 and NDTI_S1. Features based on shortwave infrared bands were very prominent across both countries; STI_S2 and STI_S3 placed second and third for Kenya, and NDTI_S2, NDTI_S1, STI_S3 and SWIR_S3 ranked fourth, fifth, eighth and ninth respectively in Tanzania. Sentinel-1 features showed less importance in this classification, with only VV_S1 in the top 10 for Kenya and RATIO_RLSPCK_S1 in the top 10 for Tanzania.

The test accuracies for the MZL task were slightly lower than for the CRL task, which was expected given that discriminating crops from each other is a harder task than discriminating crops from other land cover. Kenya achieved a 67% accuracy, while Tanzania achieved 79%. From the learning curve in Fig. 2, we can see that, as with CRL, accuracy reached a maximum at around 20 features in each country. This task relied much more heavily on the Sentinel 2 features, with large gaps in accuracy when training models only on Sentinel 1 features. Using the chosen hyperparameters, we re-trained the classifier in GEE and overlaid the maize classification onto the cropland classification, producing the maize mask in Fig. 3.

4. Maize yield mapping

4.1. Scalable crop yield mapper

Following our previous studies (Lobell et al., 2015; Azzari et al., 2017; Burke and Lobell, 2017; Jin et al., 2017b), the scalable crop yield mapper (SCYM) approach was used to estimate maize yields. Briefly, SCYM is a tool that translates satellite-observed vegetation index (VI) and gridded weather product into the maize yields using a regional-specific regression model. Unlike many other empirical methods (Bolton and Friedl, 2013; Johnson, 2014) that requires ground measurements or statistics to train their predictive model, the SCYM approach utilizes pseudo training data of thousands of paired seasonal phenology curves and yields that are simulated by a well-validated crop model (i.e. the Agricultural Production Systems Simulator (APSIM)-Maize model, Holzworth et al., 2014), spanning a realistic range of environmental and management scenarios.

Compared with previous versions of SCYM, this study improved the approach in at least two aspects. First, the crop model simulation was implemented in a more generalized way such that the simulation protocol and input dataset are ready to be extended to an entire region or continent. Second, instead of using a two-windows approach where only the greenest VI from an early and a late growing season period was fed into the predictive regression model (Lobell et al., 2015), a parametric fitting approach was proposed to derive the peak VI for training regression models. This approach better summarized the geometric characteristics of both the simulated and satellite-observed crop VI curves and thus allowed more information from the high temporal coverage of Sentinel-2 to be ingested into the estimation of crop yields.

4.1.1. Training dataset from APSIM simulations

We divided the study area into several zones of relatively homogeneous agroecological conditions, and generated separate sets of APSIM simulations for each zone. The delineation of agroecological zones was based on the Global Agro-Environmental Stratification (GAES) map (Mücher et al., 2016), which was produced by applying segmentation techniques to datasets that comprehensively described a region's agro-environmental characteristics, including climatic regimes, soil, terrain, vegetation, water availability and land cover properties. We dropped zones where < 1% of the area was classified as maize, and we combined two small zones in west Kenya to ensure enough weather

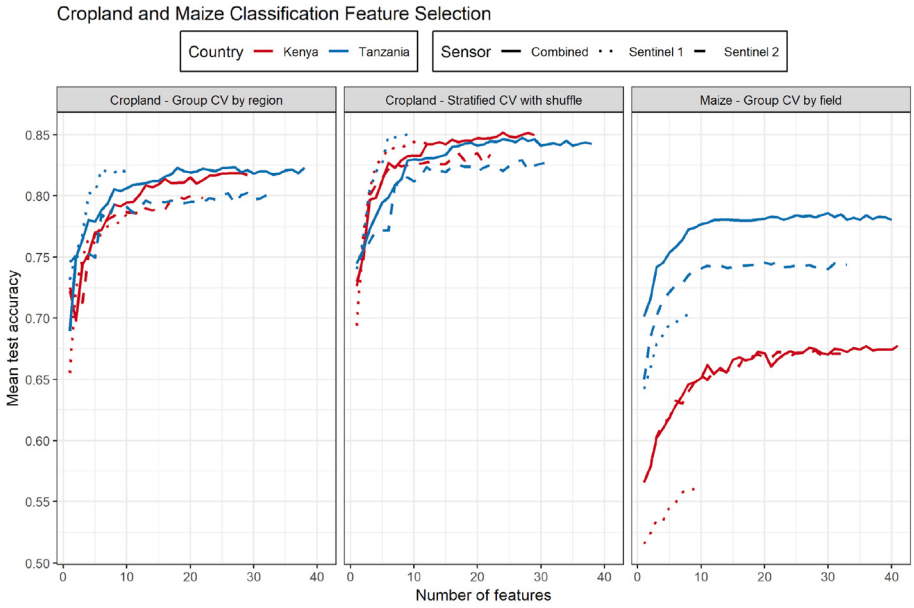

Fig. 2. Cross-validation learning curve for cropland and maize classification tasks.

gradient within zones (see Section 6 for discussions). To select the locations for running APSIM model, we first split all maize pixels into seven septiles based on the mean annual MODIS land surface temperature (which is detailed in Section 4.1.3), and randomly sampled one pixel from each septile. The GAES zones are given in Fig. 5a. Weather inputs for the period of 2009 to 2017, including maximum and minimum temperature, precipitation and solar radiation, were collected through the aWhere daily observed weather API. aWhere is an agronomic weather data platform that provides localized weather data by integrating various sources of ground stations and weather satellites. Effective data resolution is roughly 0.1 deg for sub-Saharan African countries. Soil variables were derived from the Harvest Choice 27 (HC27) generic soil profiles (Koo and Dimes, 2013). This 5 arc-minute (~10 km) gridded product was generated by clustering the 10-km grids of Harmonized World Soil Database (HWSD) v1.1 based on three criteria that crop model simulated yields are most sensitive to, including: texture, rooting depth (proxy of water availability), and organic carbon content (proxy of fertility). Maize cultivars were assigned according to the pAPSIM cultivar map, which is a composite of 10 modern hybrids and 10 open pollinated varieties with different growing season length (Elliott et al., 2013). Maize sowing dates were derived by varying the 1-km FAO's crop calendar (Van Hoolst et al., 2016). In this product, the start and end of the season were retrieved by applying a threshold approach to the 3-year NDVI profile of land pixels, which can better capture growing seasons that crossed calendar years and separate primary and secondary season. Although this crop calendar was not specific to maize, it should be adequate for the simulation purpose because our random sites were located in maize-dominant regions. We only considered the primary season, and varied the mean sowing date by ± 15 days to mimic farmers sowing practice. In addition, we varied the initial soil moisture (i.e. 60%, 80% and 100% of the extractable soil water), seeding rates (i.e. 3, 4, 5 plants/m^2), and fertilizer rates (i.e. 0, 25, 50 kg/ha N). The total number of simulations across all site × year × management combinations for training was approximately 50,000.

It worth noting that APSIM simulates canopy N, which is not

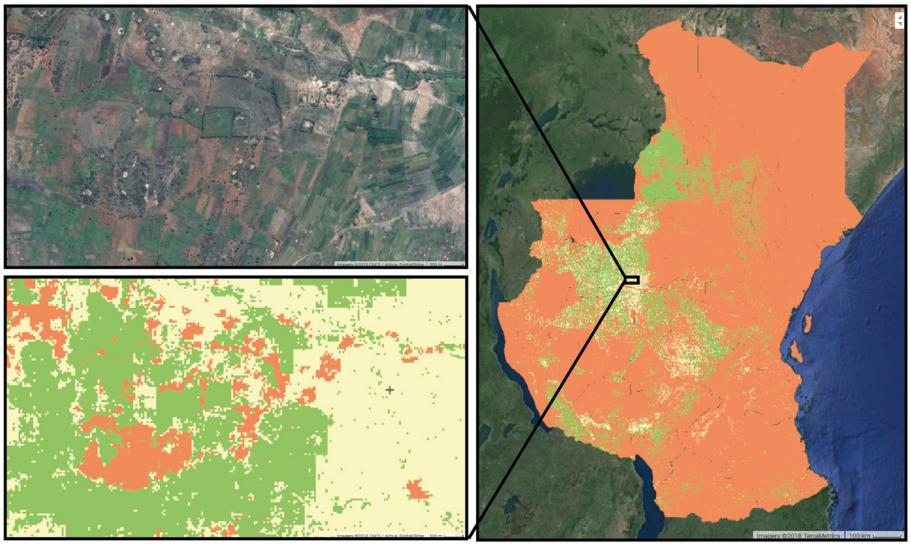

Fig. 3. Classification map with non-crop in orange, non-maize crop in yellow, and maize in green. Left detail panels show DigitalGlobe RGB imagery with the classification for comparison. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

directly observed by satellite. Therefore, we further converted canopy N to the Green Chlorophyll Vegetation Index (GCVI) based on an empirical relationship derived from a large dataset of field measurements (Schlemmer et al., 2013):

$$GCVI = 0.8 \times Canopy\ N \quad (1)$$

Compared with other vegetation indices that typically build upon red and near-infrared bands, GCVI utilize green reflectance which is more sensitive to canopy nitrogen, an important factor in Africa's low nutrient farming system (Burke and Lobell, 2017; Jin et al., 2017a).

Finally, rather than using the simulated yield, we used the simulated biomass multiplied by a constant harvest index (HI) to derive the yields for training the regression model, which has given superior performance than directly using the model simulated yields (Jin et al., 2017a, 2017b). In this study, a constant HI was prescribed as 0.4 for all GAES zones. Spatial variations in HI are not considered due to lack of geospatially explicit data on varieties and management practices that collectively determine the HI. We used a lower HI value than the typical value of 0.5 in the high yielding US Corn-Belt because the use of high yielding, and hence high HI, hybrids is very limited in Kenya and Tanzania.

4.1.2. Harmonics derived peak VI

The full time series of cloud-free Sentinel-2 GCVI observations per pixel per growing season were used to fit a harmonic regression, namely the discrete Fourier transform (DFT). As noted by Moody and Johnson (2001), the DFT can concisely summarize the seasonal pattern of land surface phenology by simply using its first and second harmonics. Here, we decomposed GCVI as a time-dependent function $f(t)$ and performed a harmonic regression as:

$$f(t) = c + \sum_{k=1}^{n} [a_k \cos(2\pi\omega kt) + b_k \sin(2\pi\omega kt)] \quad (2)$$

where a_k, b_k and c are cosine, sine and the intercept coefficients, respectively. t is the time domain in simulations ranging from 60 days before to 180 days after sowing. We set n equal to 2 following Moody and Johnson (2001). Although higher order of n was able to generate a more flexible curve that can capture local fluctuation, it was prone to overfitting for our study years during which the number of cloud-free S2 images were still limited (Fig. S5). The value of ω was set to 1.2 to match the typical seasonality in East Africa. Regression coefficients for the harmonics were estimated using the least-square error method implemented by the Google Earth Engine.

4.1.3. Yield estimation

In this study, maize yields were estimated using the following model. First, we fit both simulated yields and simulated GCVI (specifically, the peak value of the harmonic fit to simulated GCVI) to temperature and precipitation to exclude the confounding weather components:

$$Yield = \beta_{w0} + \beta_{w1} \times T + \beta_{w1} \times T^2 + \beta_{w3} \times P + \beta_{w4} \times P^2 + \varepsilon_w \quad (3)$$

$$GCVI = \beta_{v0} + \beta_{v1} \times T + \beta_{v1} \times T^2 + \beta_{v3} \times P + \beta_{v4} \times P^2 + \varepsilon_v \quad (4)$$

where T and P were the mean temperature and total precipitation during the sensitive growing season window. Then we build a regression for the two residuals, which quantified the amount of yield variation explained by GCVI variations:

$$\varepsilon_w = \beta_0 + \beta_1 \times \varepsilon_v \quad (5)$$

Finally, the yield was predicted as:

$$\widehat{Yield} = \beta_{w0} + \beta_{w1} \times T + \beta_{w1} \times T^2 + \beta_{w3} \times P + \beta_{w4} \times P^2 + \beta_0 + \beta_1 \times [GCVI - \beta_{v0} + \beta_{v1} \times T + \beta_{v1} \times T^2 + \beta_{v3} \times P + \beta_{v4} \times P^2] \quad (6)$$

The two-step modeling approach is mainly designed for addressing multicollinearity among predictive variables. Specifically, the strong colinearity between weather and peak GCVI in the simulations occasionally led to unstable coefficients and poor performance in some zones when using a single-step model, The model was built separately for each agroecological zone, with the zonal-specific sensitive window defined as 30 to 120 days after the zonal-averaged sowing date. With this regression model, we tested two strategies of sampling: one with stratified sampling as described above (hereafter the general approach), and one with localized site sampling (hereafter the local approach). The local approach incorporates prior knowledge (e.g. weather and yield gradient) about the local area when selecting simulation sites, and was only tested in west Kenya region where we collected crop cuts.

The model above uses only the fitted peak GCVI from the harmonic fit to the simulations, which risks removing valuable information on other aspects of canopy growth dynamics. However, adding the phase and amplitude only slightly increased the model fit and led to instability in predictions because of the frequently high correlation between the peak, phase, and magnitude of the harmonic fits. Therefore, the final yield regressions relied only on peak GCVI from the harmonics.

With these regression models, yield estimations at pixel level were implemented on GEE. The input peak VI was derived from harmonic regression with S2 GCVI as described in Section 4.1.2. Seasonal mean temperature was calculated with the 1 km MODerate-resolution Imaging Spectroradiometer (MODIS) Aqua MYD11A2 (8-day composite) version 6 land surface temperature (LST) product. A 2 × 2 mean filter was applied to the raw LST product to reduce the occurrence of abnormal values. Seasonal total rainfall was derived from the 0.05 arc degree Climate Hazards Group InfraRed Precipitation (CHIRPS) version 2 daily product (Funk et al., 2015). Both Aqua LST and CHIRPS are available in GEE and thus can be easily accessed as inputs to the regression model.

4.2. TOA vs SR effects on yield estimates

Several existing studies shown that using atmospherically corrected surface reflectance composites (also referred to L2A products) often had better performance over the TOA products in terms of multi-temporal analysis (e.g. land surface phenology) and the accurate retrieval of vegetation indices (Vuolo et al., 2016; Shelestov et al., 2017). We thus hypothesized that using Sentinel-2 SR composites can outperform TOA composites in the predicted yield spread. As discussed above, however, neither standard Sentinel-2 SR product nor a widely acceptable on-the-fly algorithm that converts the TOA product to the SR version is available on GEE.

To test our hypothesis, we compared three sets of yield estimates in the west Kenya region using: (i) raw Sentinel-2 TOA composites derived GCVI (TOA-GCVI) and the subsequent yields (TOA-yields), (ii) Sen2Cor derived SR GCVI (SR-L2A-GCVI) and the subsequent yields (SR-L2A-yields), and (iii) GCVI (SR-linear-GCVI) and hence yields (SR-linear-yields) derived from linearly corrected TOA composites.

Sen2Cor (current version 2.5.5) is a third-party program that perform the atmospheric-, terrain and cirrus correction of Sentinel-2 TOA L2C input data, and creates SR images. The correction is based on look-up-tables (LUTs), which were pre-calculated using a radiative transfer routine that considered different conditions of aerosols, atmosphere, ozone concentration and water vapor columns (Müller-Wilm et al., 2016). All Sentinel-2 imagery from 03/01/2017 to 10/31/2017 with < 80% cloud were processed with Sen2Cor for the west Kenya region where ground-based yield data were collected.

For comparison with the computationally intensive Sen2Cor approach, we also tested a linear correction based on the following formula:

$$\text{SR-linear-GCVI} = \beta 0 + \beta 1 \times \text{TOA-linear-GCVI} \quad (7)$$

where the coefficients β0 and β1 were derived from 1000 pairs of cropland pixels randomly sampled from the TOA-GCVI and SR-L2A-

GCVI imagery from 03/01/2017 to 10/31/2017 with < 80% cloud. The slope, β1, varied only between 1.71 and 1.97 during the growing season for the west Kenya region (Fig. S7), suggesting a fairly consistent relationship between the TOA and SR GCVI for the study area and period in this paper and hence the potential of a simple linear correction. However, researchers who want to apply this approach in other country should explore localized relationships first because atmospheric and illumination conditions may differ substantially. All data together, SR-L1A-GCVI and TOA-GCVI are highly correlated (SR-L1C-GCVI = 1.86 × TOA-linear-GCVI − 0.14, $R^2 = 0.91$). Moreover, we tested the spatial generality of β0 and β1 over the space, and found that the variations were small along a 19-tiles gradient from the most northern part of Kenya to the most southern part of Tanzania (Fig. S6). We thus combined all the 19 tiles to derive a generic equation (Fig. S8), and applied this equation to generate SR-linear-GCVI on GEE for the entire study area, although we focused on west Kenya for comparing yield estimates.

4.3. SCYM validation

Compared with aggregated district level yields from 1AF crop cuts in west Kenya, SCYM estimates were able to capture about 40–55% of the yield variation, depending on the approaches and input data (Fig. 4). The local approach outperformed the general approach in all paired cases. For 2016, using local approach increased R^2 by 25% and lowered RMSE from 0.93 t/ha to 0.89 t/ha. This level of performance is slightly better than results in Jin et al. (2017a) using the same set of crop cuts data of 2016 in this region, reflecting the advantages of using harmonic regression to retrieve geometric features of satellite observed VIs during the growing season, even though the number of cloud-free dates fed into the harmonic regression was small (i.e. 2–4 times within the growing season and 3–6 times during the non-growing season). For 2017, during which the image availability is much better (Fig. S5), SCYM yield estimates outperformed those for 2016 with both general and local approach (Fig. 6a–d). Using SR versus TOA GCVI did not always improve R^2 but increased the yield spread and lowered the RMSE (Fig. S9). In short, both general and local approaches can be used to generate reasonable yield estimates, although the local approach with more careful selection of sites for simulations led to slightly better performance.

Interestingly, even when assuming a low HI of 0.4, yield estimates based on SR product tended to slightly overpredict crop cut yields. This likely results from the fact that SCYM is calibrated using APSIM model simulations, and these simulations may not fully account for yield losses from various factors (e.g. pests or diseases) or may underestimate leaf area associated with a given yield. This can be potentially solved by calibrating the phenology module in the crop model, as was shown in Jin et al. (2017b) for maize in the US Midwest, but was outside the scope of this study.

4.4. Yield maps for years 2016–2018

Fig. 5 displays the maize yield estimates for 2017 for the entire study region, as well as yields over the 2016–2018 period for a subset of the region. Predicted yields in the study area were higher in west Kenya, and decreased southward towards the border with Tanzania, and then slightly increased again towards south Tanzania (Fig. 5a). Yields can be as high as 5 t/ha in west Kenya, but were typically around 2–3 t/ha in other parts of Kenya. Importantly, SCYM estimates identified substantial fine-scale heterogeneity in yields, even in regions where average yields were low (Fig. 5c). The map also revealed some interesting discontinuity across administrative boundaries. For instance, nearby the boundary of Kakamega and Nandi county, fields were obviously smaller in size and less productive on the Kakamega side compared to the Nandi side (Fig. 5b). The annual maps shown in Fig. 5d suggest that spatial yield gradients were similar among different years, but with subtle differences. Overall, the rich variability in these maps indicates the combined effects of myriad factors on maize productivity. To further explore some of these factors, and further illustrate the value of GEE and the yield mapping approach presented here, the following section discusses two specific applications of these yield maps.

5. Applications of yield maps

We present two analytical examples of exploring yield heterogeneity and soil constraints focused on west Kenya. We focus on this region because: (i) satellite imagery was better than other regions in terms of quality and quantity (Fig. S5), which led to greater reliability in yield estimates (Fig. 6d); and (ii) this 40,000 km^2 is a core production area in Kenya and exhibits substantial yield heterogeneity over the landscape corresponding to a weather gradient (Fig. S10).

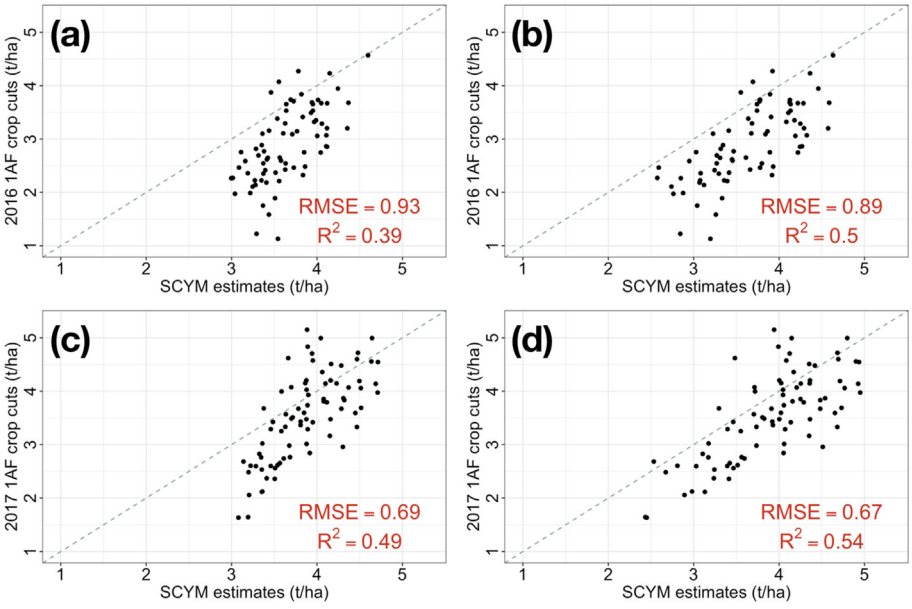

Fig. 4. Comparisons of SCYM yield estimates in west Kenya with crop cuts aggregated to district level in (a–b) 2016 and (c–d) 2017 main maize season. Left panels (a, c) show results using the "general" model trained on simulations from sites randomly selected throughout the GAES zone, while right panels (b, d) show results with "local" model using only sites from within the region with ground data. Importantly, for all panels none of the 1AF data was used to calibrate the simulation.

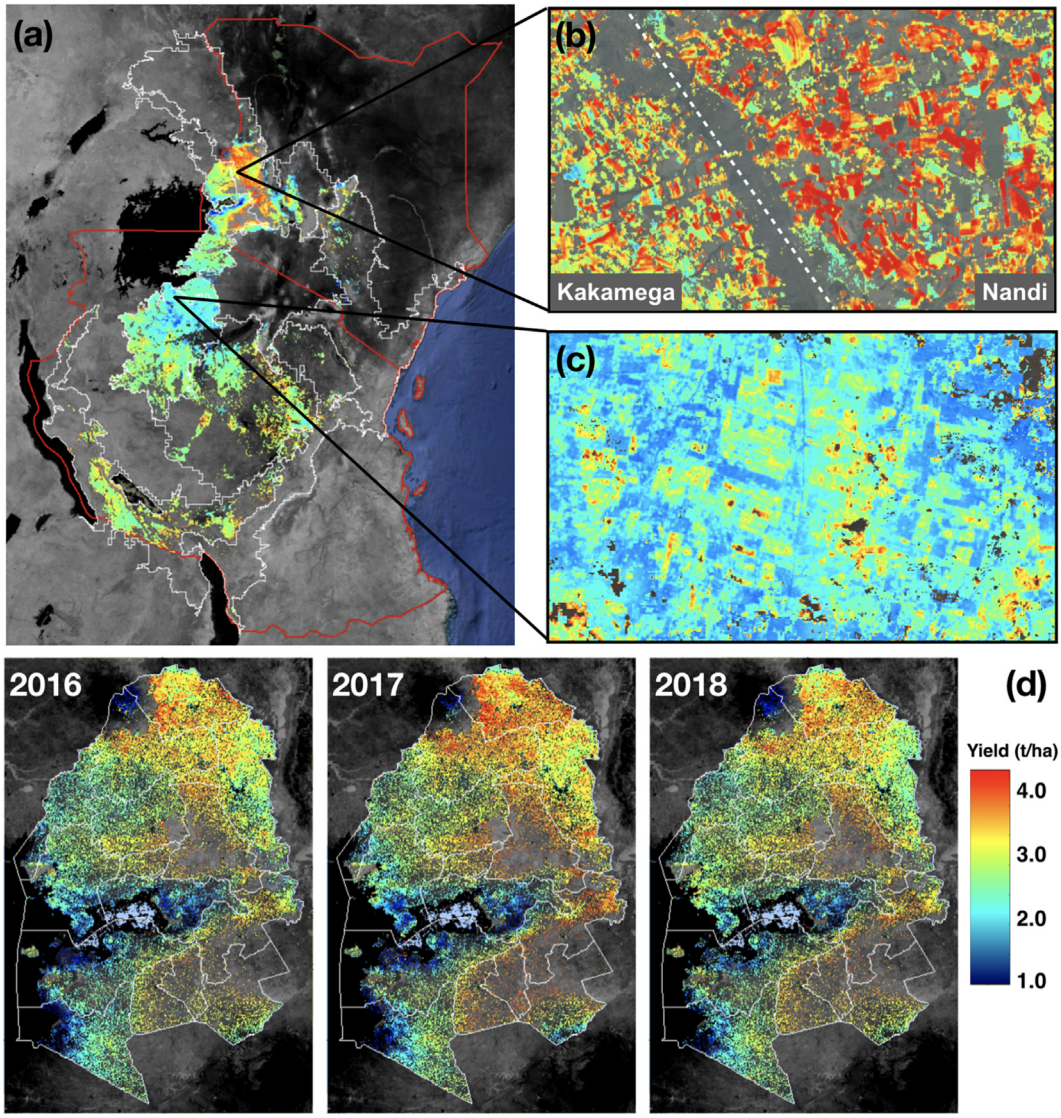

Fig. 5. Study region of Kenya and Tanzania showing (a) estimated maize yields for 2017, including detail for two sub-regions (b and c) as well as (d) yields for 2016–18 for 15 counties in Western Kenya. White lines in (a) delineate the boundary of GAES zones in which maize growing > 1% of the area. Dashed white line in (b) is roughly the county boundary.

5.1. Yield heterogeneity analysis

Quantifying yield heterogeneity across a region can be useful in many aspects, for example, understanding yield progress and drivers of yield gaps (Lobell, 2013), optimizing management practice and designing crop insurance programs (Cooper et al., 2012). For example, Lobell and Azzari (2017) examined maize yields for 2000–2015 in the US Midwest using SCYM approach, and found both between and within fields yield heterogeneity was rising, reflecting the recent adoption of precision agriculture technology and greater sowing density in maize that disproportionately enhance yields on better soils. With a similar yield estimation approach and Random Forest analysis, Jain et al. (2017) identified later sowing and warmer temperatures were often associated with low wheat yield in India's Wheat Belt. To our knowledge, satellite-based investigation of yield heterogeneity that spanned a large region has not been done for the smallholder farming system in Africa, although efforts with a combination of crop modeling and UAV sensing exist for a much smaller scale (e.g. Tittonell et al., 2007; Schut et al., 2018).

Here, based on the yield estimates derived with the local approach (Fig. 4d), we calculated the overall district yield heterogeneity (YHo) as the percentage difference between the 95th percentile (Y95) and mean (Ymean) for each district during 2016–2018:

$$YH_o = (Y_{95} - Y_{mean})/Y_{95} \quad (8)$$

This equation was adapted from Lobell and Azzari (2017) who used it to study maize yield heterogeneity in the US Midwest. We did not calculate the within-field yield heterogeneity because boundaries of individual fields were hard to obtain at this moment. It worth noting that this derived YHo should not be interpreted as yield gaps because even the best current field performance may still be far behind the attainable yields in this typical low-input sub-Saharan Africa region (Sheahan and Barrett, 2017). In addition, we sampled ca. 30–200 pixels per district-year in proportion to the number of maize pixels to estimate the full yield distribution for this region. Finally, to understand how consistent patterns of spatial variability are from year to year, for each district we calculated the Pearson correlation for all pixel-level (ranging from 66,038 to 6,963,705 pixels) yields of 2016 and 2017 using GEE. Persistence of spatial yield variability has been used as an indicator of how much location-specific (e.g. soil quality, farmer skill) vs. idiosyncratic factors influence yield variability (Lobell, 2013; Farmaha et al., 2016).

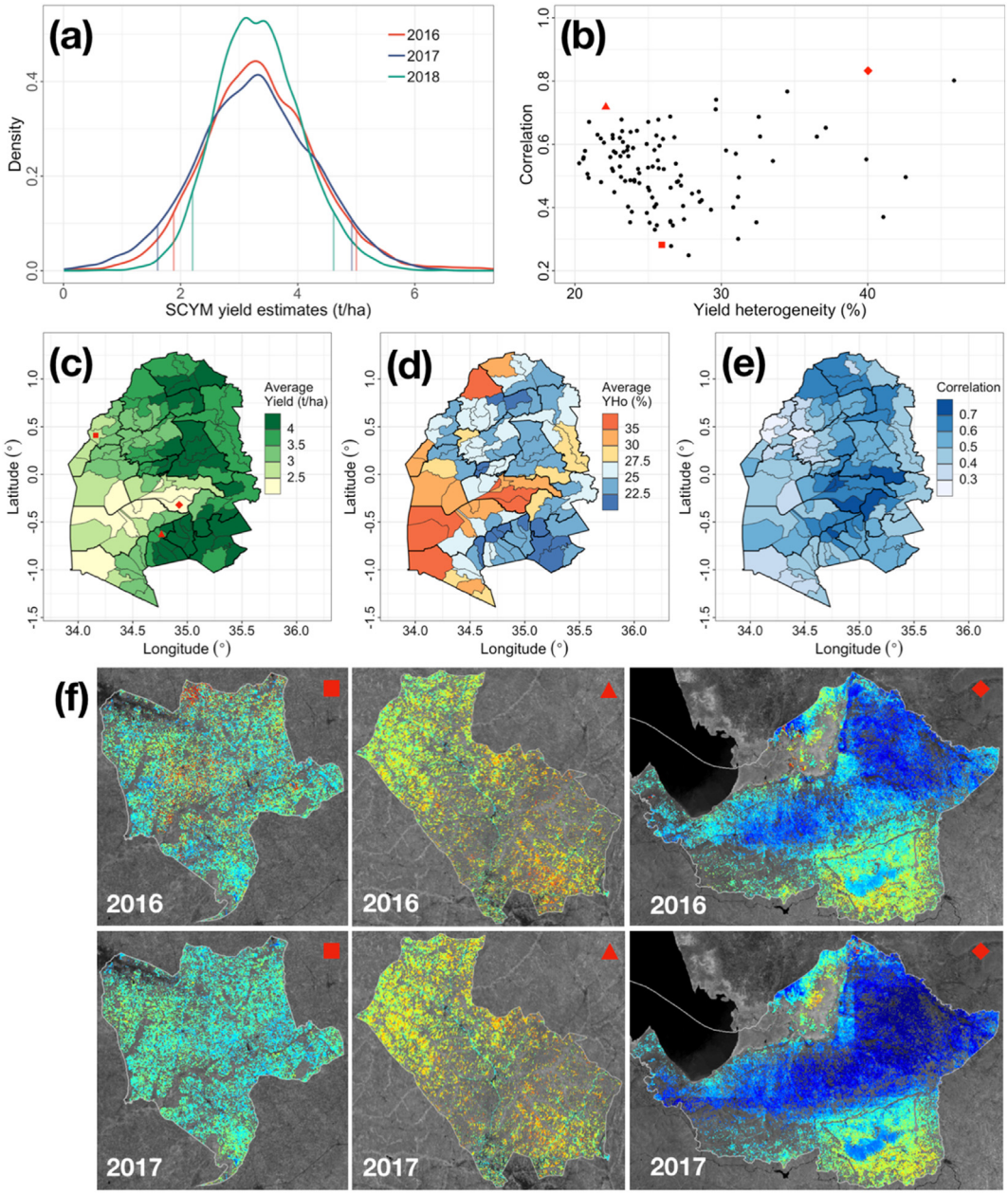

Fig. 6. Application of SCYM estimates to yield heterogeneity analysis. (a) SCYM estimated yield distribution for 2016–2018, with vertical segments indicate 5% and 95% percentile values. (b) Scatter plot of district level yield heterogeneity (YHo) and yield correlation between 2016 and 2017, with red square, diamond and triangle representing three selected districts shown in (f). (c) Aggregated district level average yield for 2017. (d) Average district level YHo for 2016 and 2017. (e) Average district level yield correlation between 2016 and 2017. (f) Yield map for three selected district in 2016 and 2017. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6a shows the histogram density of yields for the focus region for the three years. The density curves of 2016 and 2017 were quite similar, although the latter had a fatter tail at the lower end, indicating possibly more yield penalty due to weather stresses. The 2018 distribution was narrower than the other two years, indicating less yield variability. We caution that one possible explanation is that variation during the grain-filling stage was less well captured without available Sentinel-2 imagery after September by the time this paper was written, and therefore we focus subsequent analysis on 2016–17.

Values of YHo ranged from roughly 20 to 40% of mean yield, indicating that if all farmers achieved the yields at the 95th percentile of the current distribution, that yields would be increased by an average of ~30% (Fig. 6b). At the same time, it was clear that the same fields were not the highest performing across both years, with yield correlations between 2016 and 2017 mostly between 0.3 and 0.8 (Fig. 6b). Spatially, higher YHo tended to coincide with districts of lower Ymean (Fig. 6c, d). To illustrate different combinations of yield correlation and YHo, we display three representative districts (Fig. 6b, f), including Matayos from Busia county for median YHo and low yield correlation (square), Kitutu Chache South from Kitutu province for Kisii county for low YHo and high yield correlation (triangle), and Nyakach from Kisumu county for high YHo and high yield correlation (diamond). These results provided a preliminary screening of the yield heterogeneity in this region, although further data collection and study is needed in order to identify drivers of the spatial pattern as well as changes over time.

5.2. Soil constraint analysis

To understand the edaphic drivers of yield heterogeneity and help guide nutrient-specific interventions, we compared SCYM yield estimates with existing continent-wide soil hydraulic and nutrient data. We selected 11 variables from the 250 m soil nutrient maps of Sub-Saharan Africa (Hengl et al., 2017), including: organic carbon (OC), total organic nitrogen (tot_N), total phosphorus (tot_P), and extractable- potassium (K), calcium (Ca), magnesium (Mg), sodium (Na), iron (Fe), Manganese (Mn), Copper (Cu), Aluminum (Al) for 0–30 cm depth. Briefly, this product is generated using an ensemble of two machine learning algorithms with extensive training data of soil samples from > 50,000 locations and a large stack of GIS layers that proximate soil forming processes (e.g. climate, topography, lithology and vegetation) (Hengl et al., 2015). Among all the variables being predicted, the ones we selected all had cross-validation accuracy (i.e. R^2) > 0.5 (between 0.53 and 0.86). In addition, we obtained two soil hydraulic variables, root zone depth (ERZD) and root zone total plant available water holding capacity (mm) of the whole earth fraction (AWC), from the Root Zone Plant-Available Water Holding Capacity of Sub-Saharan Africa soils, version 1.0 dataset (Leenaars et al., 2018). This 1 km spatial resolution product was generated using digital soil mapping techniques by ingesting 28,000 soil profiles from SSA. Predicted volumetric moisture content (VMC) at saturation explained about 72% of the variation in observed values over 6 depth intervals, with RMSE equal to 0.102 cm^3/cm^3 (Leenaars et al., 2018). Overall, these two soil datasets represent to our knowledge the best existing maps of soil properties in the region.

We randomly sampled the stack of soil variables and SCYM estimates at 50 locations per district within the focus region, conditioning on that locations fell within maize pixels. Then we conducted correlation analysis as well as Random Forest regression (Breiman, 2001) to quantify the relative importance of these soil variables in explaining the variation of mean yield of 2016–2018. We used RF regression because it doesn't require an explicit model structure, which was attractive in this case given the potential non-linear relationship between response and explanatory variables as well as the interactions among explanatory variables. In this study, the RF algorithm created 300 different decision trees using random subsamples of the data and computed the prediction error on the out-of-bag portion of these bootstrapped datasets. For each node, 4 of all available explanatory variables were used to split the data. To calculate variance explained, this method considers the average variance explained in the out of bag sample across all trees.

Results showed a positive correlation between Al, OC, tot_N, tot_P and yields, whereas soil nutrients showed weak negative correlation with yields (Fig. 7a). The RF model was able to explain > 70% of the yield variation (Fig. 7b). A large fraction of the explained variation is due to nitrogen, with tot_N, an indicator of direct available N source, ranked the first, followed by OC, which directly reflects the amount of soil organic matter and potential mineral N that can be released through mineralization. Al, an indicator of soil acidity (which may limit plant growth) ranked high among the variables, but, surprisingly, was positively correlated with yields. This may be explained by three facts: (i) the average soil pH in this region is > 5.5, meaning acidity stress is likely uncommon; (ii) slightly acidic soils facilitate the release of yield limiting nutrients; and (iii) higher Al concentrations coincide with greater leaching and higher precipitation, hence slightly better water supply for maize growth. Interestingly, AWC and ERZD were not as important as expected, likely because the coarser 1 km resolution is not ideal for comparing with SCYM estimates at 10 m resolution. Using all variables, the predictive model did a fairly good job (R^2 = 0.72) for the 20% leave-out test dataset (Fig. 7c). Overall, our analysis confirmed the importance of indigenous N supply in this typically low-fertilizer-use region and the potential of closing yield gaps by increasing N application.

6. Discussion

The GEE platform, with its vast stores of satellite data, computational power, and implementation of various machine learning algorithms and flexible user interfaces, was central to this study. For example, GEE provided the possibility to implement a scalable solution to most of Sentinel-2 limitations. Not only we could easily sample and export contaminated pixels from a curated collection of imagery, but we were also able to re-use a light-weight decision tree trained offline to classify clouds, haze, and shadows on the fly. Similarly, applying correction coefficients to convert hundreds of images from TOA into SR observations was seamless and fast. Furthermore, the UI API allowed us to add custom interactivity to the Code Editor and transform it into a useful tool to collect labels to train the CRL model. GEE's scalability was also used to generate pixel-level multi-seasonal median composites, calculate harmonic regressions, and estimate and apply Random Forest models. At the same time, this innovative platform was not without some challenges, such as the long time needed to pre-compute and store the median composites as assets, a step needed to avoid exceeding memory limits.

The limited availability of accurate ground data on maize locations and yields precluded a systematic evaluation of our estimates across the diversity of all maize systems in Tanzania and Kenya. Nonetheless, the out-of-sample accuracies for cropland and maize land identification was

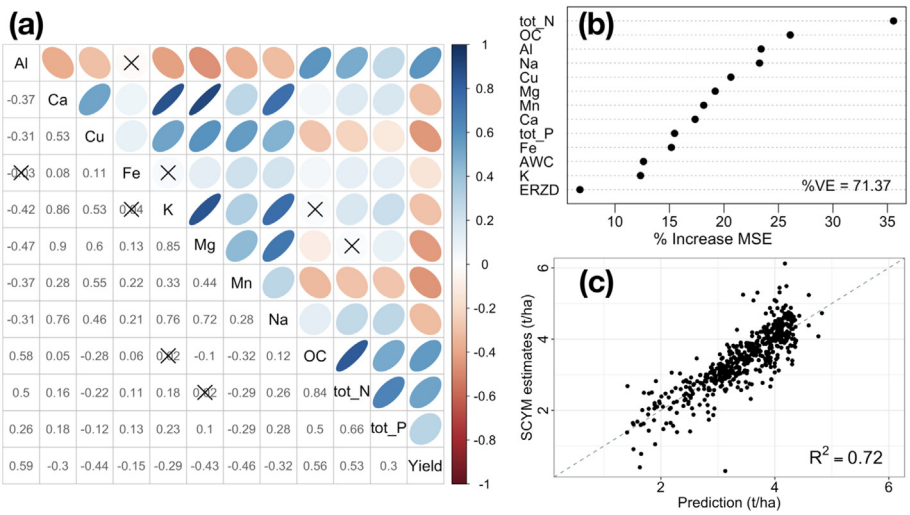

Fig. 7. Application of SCYM estimates to soil property analysis. (a) Correlation matrix plot of SCYM yields and 11 soil nutrient variables at randomly selected locations. (b) The relative importance of predictive variables in the Random Forest (RF) model, with the most important variable having the largest % increase in mean squared error (MSE) between the full model and one where that variable is permuted. (c) Prediction performance of the RF model on the 20% leave-out validation dataset.

encouraging, as was the fact that the SCYM yield estimates explained roughly half of the variation in yields in west Kenya as estimated by crop cuts. This latter result of course means that roughly half of the variation in ground-based yields was not explained by our estimates, and for some applications this performance may be viewed as insufficient. However, it is worth emphasizing that at least some fraction of the remaining variance not explained by SCYM arises because of measurement error in ground data. Prior work has shown, for instance, that yields estimated from satellites can often match or exceed ground-based estimates in terms of correlations with suspected yield constraints, such as fertilizer or hybrid seed use (Burke and Lobell, 2017). In the current study, the fact that 70% of SCYM yields could be explained by soil factors indicates that we are indeed capturing meaningful variation in yields.

The limited amount of ground data for validation also highlights the unique potential of remote sensing for agricultural assessment. For regions without ground data to enable a quantitative metric of performance, a meticulous qualitative analysis of the resulting classification may provide additional information on the performance of the underlying model. In our case, the picture that emerged from a qualitative analysis of both the CRL and MZL maps was that our models performed very well at detecting the overall regions where crop, and more specifically maize, was grown. A chain process of creating MZL map on top of the CRL map may pass some of the uncertainty in cropland classification. However, given the high complexity of land cover in our study region and the size of ground labels (although exceptionally large already), skipping the step of CRL maps will lead to worse performance in maize classification.

A major challenge addressed in this work is how to best generate pseudo-training data with crop model simulations across broad regions. This critical step in the SCYM framework determines the regression model and ultimately the performance of yield estimates. A good set of simulations should span a realistic range of variation (or spread) in both GCVI and yields. The improvement achieved by switching to more localized site selection as shown in Fig. 4c, d reflects the importance of site selection. Nonetheless, the general approach suggested here whereby sites are randomly chosen within GAES zones along a temperature gradient produced results that agreed well with ground data and produced qualitatively reasonable patterns of yields across the region (i.e. very few extremely high or near-zero yield estimates). Some remaining challenges for future work include (i) the optimal way to define zones within which sites are selected, (ii) how to fully cover the range of temperature and precipitation gradients without excessive amounts of simulation sites, and (iii) how to refine the cultivar representations (mostly parameters related to phenology and harvest index), possibly through building detailed crop sowing and harvest date maps based on improved maize classification maps and high resolution and frequent satellite observations.

Both Sentinel-1 and Sentinel-2 played key roles in this study, with the former being particularly important for crop and maize land classification and the latter for both maize classification and yield estimation. Sentinel-2 data provides global coverage with a 5 days revisit time at 10 m resolution. In addition to near-infrared and shortwave infrared bands, it has three red-edge bands, which proved useful in crop classification. This combination of factors is unprecedented for a public dataset, making Sentinel-2 a much-needed addition to the current roster of open-access satellites. At the same time, the lack of a readily available standard SR product with an accurate QA band for clouds, haze, shadows, snow, and other contaminated pixels makes the use of this dataset challenging at large scales. As the amount of satellite data increases, the availability of “analysis-ready” data will be key a factor in determining the role played by different datasets in enabling new applications and scales.

7. Conclusion

This study presented a first-of-its-kind mapping of maize yields at 10 m resolution across the entirety of two large maize producing countries, by leveraging the power of newly available satellite sensors, the GEE platform, and advanced classification and yield mapping algorithms. Accomplishing such a mapping in East Africa is extremely challenging because of the heterogeneous smallholder farming landscapes, pervasive and year-round cloud coverage, and the limited and sparse ground measurements. By using a large number of training labels, Random Forest-based classifier can successfully identify cropland and maize pixels. We note the utility of backscatter metrics in classification, particularly when cloud-free optical imagery is limited. The yield estimation approach presented in this study was able to capture about half of the variation in the crop-cut yields at the district level in Western Kenya, although multiple opportunities exist to further improve the results. Using yield maps generated in this study and other geospatial data, we were able to explore the magnitude and interannual variability of spatial heterogeneity of yields, and to investigate the soil constraints in the region. We identified soil nitrogen levels as the predominant factor, suggesting the importance of nutrient management as a mean to closing yield gaps in the region. Lessons learned in this study advanced the understanding of smallholder agriculture system in East Africa, and can be transferred to other similar system as well.

Acknowledgments

This study is made possible by the support of the American People provided to the Feed the Future Innovation Lab for Sustainable Intensification through the United States Agency for International Development (USAID). The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. Program activities are funded by the USAID under Cooperative Agreement No. AID-OAA-L-14-00006. Funding was also provided by NASA Harvest Consortium grant 54308-Z6059203 to DBL, USAID grant 740681-74171D to MB and DBL and Global Innovation Fund to MB. The authors acknowledge Jake Campolo for early assistance on the maize simulations, AfSIS, One Acre Fund and TAMASA for providing data and the Bill and Melinda Gates Foundation for supporting those efforts in data collection and dissemination. Code and data generated from this paper will be made available to public upon publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2019.04.016.

References

Abramov, S., Rubel, O., Lukin, V., Kozhemiakin, R., Kussul, N., Shelestov, A., Lavreniuk, M., 2017. Speckle reducing for Sentinel-1 SAR data. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2353–2356 (IEEE).

Azzari, G., Lobell, D., 2017. Landsat-based classification in the cloud: an opportunity for a paradigm shift in land cover monitoring. Remote Sens. Environ. 202, 64–74. https://doi.org/10.1016/j.rse.2017.05.025.

Azzari, G., Jain, M., Lobell, D.B., 2017. Towards fine resolution global maps of crop yields: testing multiple methods and satellites in three countries. Remote Sens. Environ. 202, 129–141. https://doi.org/10.1016/j.rse.2017.04.014.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84. https://doi.org/10.1016/j.agrformet.2013.01.007.

Breiman, L., 2001. Machine learning. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/a:1010933404324.

Burke, M., Lobell, D.B., 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. Proc. Natl. Acad. Sci. 114 (9), 2189–2194. https://doi.org/10.1073/pnas.1616919114.

Carletto, C., Jolliffe, D., Banerjee, R., 2015. From tragedy to renaissance: improving agricultural data for better policies. J. Dev. Stud. 51 (2), 133–148. https://doi.org/10.1080/00220388.2014.968140.
Chang, J., Hansen, M.C., Pittman, K., Carroll, M., DiMiceli, C., 2007. Corn and soybean mapping in the United States using MODIS time-series data sets. Agron. J. 99 (6), 1654–1664. http://doi:10.2134/agronj2007.0170.
Chivasa, W., Mutanga, O., Biradar, C., 2017. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: a review. Int. J. Remote Sens. 38 (23), 6816–6845. https://doi.org/10.1080/01431161.2017.1365390.
Cooper, J., Zulauf, C., Langemeier, M., Schnitkey, G., 2012. Implications of within county yield heterogeneity for modeling crop insurance premiums. Agric. Finance Rev. 72 (1), 134–155. https://doi.org/10.1108/00021461211222213.
Dasari, K., Anjaneyulu, L., Jayasri, P., Prasad, A., 2015. Importance of speckle filtering in image classification of SAR data. In: 2015 International Conference on Microwave, Optical and Communication Engineering (ICMOCE), https://doi.org/10.1109/icmoce.2015.7489764.
Elliott, J., Kelly, D., Best, N., Wilde, M., Glotter, M., Foster, I., 2013. The parallel system for integrating impact models and sectors (pSIMS). In: Proceedings of the Conference on Extreme Science and Engineering Discovery Environment Gateway to Discovery - XSEDE '13, https://doi.org/10.1145/2484762.2484814.
Farmaha, B.S., Lobell, D.B., Boone, K.E., Cassman, K.G., Yang, H.S., Grassini, P., 2016. Contribution of persistent factors to yield gaps in high-yield irrigated maize. Field Crop Res. 186, 124–132. https://doi.org/10.1016/j.fcr.2015.10.020.
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66.
Food and Agriculture Organization of the United Nations, 2018. FAOSTAT Database. FAO, Rome, Italy(Retrieved Mar 15th, 2019 from http://www.fao.org/faostat/en/#data/QC).
Goodman, J.W., 1976. Some fundamental properties of speckle. J. Opt. Soc. Am. 66 (11), 1145. https://doi.org/10.1364/josa.66.001145.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.
Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Shepherd, K.D., ... Tondoh, J.E., 2015. Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. PLoS One 10 (6). https://doi.org/10.1371/journal.pone.0125814.
Hengl, T., Leenaars, J.G., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B., Mamo, T., ... Wheeler, I., 2017. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr. Cycl. Agroecosyst. 109 (1), 77–102.
Hollstein, A., Segl, K., Guanter, L., Brell, M., Enesco, M., 2016. Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. Remote Sens. 8 (8), 666. https://doi.org/10.3390/rs8080666.
Holzworth, D.P., Huth, N.I., Devoil, P.G., Zurcher, E.J., Herrmann, N.I., Mclean, G., ... Keating, B.A., 2014. APSIM – evolution towards a new generation of agricultural systems simulation. Environ. Model Softw. 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009.
Hoolst, R.V., Eerens, H., Haesen, D., Royer, A., Bydekerke, L., Rojas, O., ... Racionzer, P., 2016. FAO's AVHRR-based Agricultural Stress Index System (ASIS) for global drought monitoring. Int. J. Remote Sens. 37 (2), 418–439. https://doi.org/10.1080/01431161.2015.1126378.
Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., ... Koetz, B., 2015. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. Remote Sens. 7 (9), 12356–12379. https://doi.org/10.3390/rs70912356.
Jain, M., Srivastava, A., Balwinder-Singh, Joon, R., Mcdonald, A., Royal, K., … Lobell, D. (2016). Mapping smallholder wheat yields and sowing dates using micro-satellite data. Remote Sens., 8(10), 860. doi:https://doi.org/10.3390/rs8100860.
Jain, M., Singh, B., Srivastava, A.A.K., Malik, R.K., Mcdonald, A.J., Lobell, D.B., 2017. Using satellite data to identify the causes of and potential solutions for yield gaps in India's Wheat Belt. Environ. Res. Lett. 12 (9), 094011. https://doi.org/10.1088/1748-9326/aa8228.
Jin, Z., Azzari, G., Burke, M., Aston, S., Lobell, D., 2017a. Mapping smallholder yield heterogeneity at multiple scales in Eastern Africa. Remote Sens. 9 (9), 931. https://doi.org/10.3390/rs9090931.
Jin, Z., Azzari, G., Lobell, D.B., 2017b. Improving the accuracy of satellite-based high-resolution yield estimation: a test of multiple scalable approaches. Agric. For. Meteorol. 247, 207–220. https://doi.org/10.1016/j.agrformet.2017.08.001.
Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116–128. https://doi.org/10.1016/j.rse.2013.10.027.
Koo, J., Dimes, J., 2013. HC27 Generic Soil Profile Database. Harvard Dataverse Ver. 4. International Food Policy Research Institute, Washington, DC.
Kraskov, A., Stögbauer, H., Grassberger, P., 2004. Estimating mutual information. Phys. Rev. E 69 (6). https://doi.org/10.1103/physreve.69.066138.
Lambert, M.-J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. Remote Sens. Environ. 216, 647–657. https://doi.org/10.1016/j.rse.2018.06.036.
Leenaars, J.G., Claessens, L., Heuvelink, G.B., Hengl, T., González, M.R., Bussel, L.G.V., ... Cassman, K.G., 2018. Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa. Geoderma 324, 18–36. https://doi.org/10.1016/j.geoderma.2018.02.046.
Lobell, D.B., 2013. The use of satellite data for crop yield gap analysis. Field Crop Res. 143, 56–64. https://doi.org/10.1016/j.fcr.2012.08.008.
Lobell, D.B., Azzari, G., 2017. Satellite detection of rising maize yield heterogeneity in the U.S. Midwest. Environ. Res. Lett. 12 (1), 014014. https://doi.org/10.1088/1748-9326/aa5371.
Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. 164, 324–333. https://doi.org/10.1016/j.rse.2015.04.021.
Lowder, S.K., Skoet, J., Raney, T., 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. World Dev. 87, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041.
Mccullough, E.B., 2017. Labor productivity and employment gaps in Sub-Saharan Africa. Food Policy 67, 133–152. https://doi.org/10.1016/j.foodpol.2016.09.013.
Moody, A., Johnson, D.M., 2001. Land-surface phenologies from AVHRR using the discrete Fourier transform. Remote Sens. Environ. 75 (3), 305–323. https://doi.org/10.1016/s0034-4257(00)00175-9.
Mücher, Sander, Simone, L.D., Kramer, H., Wit, A.D., Roupioz, L., Hazeu, G., ... Cormont, A., 2016. A New Global Agro-environmental Stratification (GAES). https://doi.org/10.18174/400815.
Müller-Wilm, U., Devignot, O., Pessiot, L., 2016. Sen2Cor configuration and user manual. https://step.esa.int/thirdparties/sen2cor/2.4.0/Sen2Cor_240_Documenation_PDF/S2-PDGS-MPC-L2A-SUM-V2.4.0.pdf.
Ross, B.C., 2014. Mutual information between discrete and continuous data sets. PLoS One 9 (2). https://doi.org/10.1371/journal.pone.0087357.
Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M., West, P.C., 2016. Subnational distribution of average farm size and smallholder contributions to global food production. Environ. Res. Lett. 11 (12), 124010. https://doi.org/10.1088/1748-9326/11/12/124010.
Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., Rundquist, D., 2013. Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. Int. J. Appl. Earth Obs. Geoinf. 25, 47–54. https://doi.org/10.1016/j.jag.2013.04.003.
Schut, A.G., Traore, P.C.S., Blaes, X., By, R.A.D., 2018. Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites. Field Crop Res. 221, 98–107. https://doi.org/10.1016/j.fcr.2018.02.018.
Sheahan, M., Barrett, C.B., 2017. Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Policy 67, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010.
Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S., 2017. Exploring Google Earth Engine platform for big data processing: classification of multi-temporal satellite imagery for crop mapping. Front. Earth Sci. 5. https://doi.org/10.3389/feart.2017.00017.
Therneau, T.M., Atkinson, E.J., 1997. An introduction to recursive partitioning using the RPART routines.
Tittonell, P., Vanlauwe, B., Ridder, N.D., Giller, K., 2007. Heterogeneity of crop productivity and resource use efficiency within smallholder Kenyan farms: soil fertility gradients or management intensity gradients? Agric. Syst. 94 (2), 376–390. https://doi.org/10.1016/j.agsy.2006.10.012.
Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., ... Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24. https://doi.org/10.1016/j.rse.2011.05.028.
Vuolo, F., Żółtak, M., Pipitone, C., Zappa, L., Wenng, H., Immitzer, M., ... Atzberger, C., 2016. Data service platform for Sentinel-2 surface reflectance and value-added products: system use and examples. Remote Sens. 8 (11), 938. https://doi.org/10.3390/rs8110938.
Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I., Justice, C.O., 2015. Cloud cover throughout the agricultural growing season: impacts on passive optical earth observations. Remote Sens. Environ. 156, 438–447. https://doi.org/10.1016/j.rse.2014.10.