June 2017

NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) @ 30-m for Africa: Cropland Extent Product (GFSAD30AFCE)

Algorithm Theoretical Basis Document (ATBD)

USGS EROS
Sioux Falls, South Dakota

Document History

Document Version	Publication Date	Description
1.0	September, 2017	Original
1.1		Modification made according to USGS reviewer's comments

Contents

I. Members of the team 4
II. Historical Context and Background Information 4
III. Rationale for Development of the Algorithms 5
IV. Algorithm Description 6
a. Input data 7
1. Reference Croplands Samples 7
2. Satellite Imagery: Sentinel-2 and Landsat-8 8
b. Theoretical description 10
1. Definition of Croplands 10
2. Algorithms 11
c. Practical description 11
1. Random Forest (RF) Algorithm 11
2. Support Vector Machines (SVMs) Algorithm 12
3. Combining Random Forest & Support Vector Machines for optimal results 12
4. Recursive Hierarchical Image Segmentation (RHSeg) 13
5. Integration of pixel-based classification and hierarchical segmentation 14
6. Programming and codes 15
7. Results 15
8. Cropland areas of Africa 16
V. Calibration Needs/Validation Activities 19
VI. Constraints and Limitations 20
VII. Publications 21
VIII. Acknowledgements 25
IX. Contact Information 25
X. Citations 26
X. References 26

I. Members of the team

This Global Food Security-support Analysis Data 30-m (GFSAD30) Cropland Extent Product of Africa (GFSAD30AFCE) was produced by the following team members. Their specific role is mentioned.

Dr. Jun Xiong, Research Scientist, Bay Area Environmental Research Institute (BAERI) at the United States Geological Survey (USGS), led the GFSAD30AFCE product generation effort. Dr. Xiong was instrumental in the design, coding, computing, analyzing, and synthesis of the Sentinel-2 and Landsat-8 derived nominal 30-m GFSAD30AFCE cropland product of the African continent for the nominal year 2015. He was also instrumental in writing the manuscripts, ATBD, and user documentation.

Dr. Prasad S. Thenkabail, Research Geographer, United States Geological Survey, is the Principal Investigator (PI) of the GFSAD30 project. Dr. Thenkabail was instrumental in developing the conceptual framework of the GFSAD30 project and the GFSAD30AFCE product. He made significant contribution in writing the manuscripts, ATBD, user documentation, and providing scientific guidance on the GFSAD30 project.

Dr. James C. Tilton, Computer Engineer with the Computational and Information Sciences and Technology Office (CISTO) of the Science and Exploration Directorate at the NASA Goddard Space Flight Center, developed the recursive hierarchical segmentation (RHSeg) algorithm. He also helped Dr. Xiong implement RHSeg at NASA's supercomputer facility.

Dr. Murali Krishna Gumma, Senior Scientist at the International Crops Research Institute for the Semi-Arid Tropics (**ICRISAT**), helped collect reference data critical for training the machine learning algorithms.

Dr. Pardhasaradhi Teluguntla, Research Scientist, Bay Area Environmental Research Institute (BAERI) at the United States Geological Survey (USGS), provided input and insights on cropland extent product generation for the African continent.

Dr. Russell G. Congalton, Professor of Remote Sensing and GIS at the University of New Hampshire, led the independent accuracy assessment of the entire GFSAD30 project including GFSAD30AFCE.

Mr. Justin Poehnelt, Computer Scientist with the United States Geological Survey , contributed to the initial conceptualization and development of the croplands.org website.

Ms. Kamini Yadav, PhD student at the University of New Hampshire, made major contributions to the independent accuracy assessment directed by Prof. Russell G. Congalton.

Mr. Richard Massey, PhD student at the Northern Arizona University, shared his expertise in cloud computing and in RHSeg algorithm's implementation on supercomputers.

II. Historical Context and Background Information

Monitoring global croplands is imperative for ensuring sustainable water and food security to the people of the world in the twenty-first century. However, the currently available cropland products suffer from major limitations such as: (1) the absence of precise spatial locations of the cropped areas; (2) coarse resolution of map products with significant uncertainties in areas, locations, and detail; (3) uncertainties in differentiating irrigated areas from rainfed areas; (4) absence of crop type information and cropping intensities; and/or (5) the absence of a dedicated Internet data portal for the dissemination of cropland products. Therefore, the Global Food Security-support Analysis Data (GFSAD) project aimed to address these limitations by producing cropland maps at 30m resolution covering the globe, referred to as Global Food Security Support-Analysis Data @ 30-m (GFSAD30)

products. This Algorithm Theoretical Basis Document (ATBD) provides a basis upon which the GFSAD30 cropland extent product was developed for the continent of Africa (GFSAD30AFCE, Table 1).

Table 1. Basic information of the Global food security support-analysis data @ 30-m cropland extent product for the African continent (GFSAD30AFCE

Product Name	Short Name	Spatial Resolution	Temporal Resolution
GFSAD 30-m Cropland Extent Product of Africa	GFSAD30AFCE	30-m	nominal 2015

III. Rationale for Development of the Algorithms

Mapping the precise location of croplands enables the extent and area of agricultural lands to be more effectively captured, which is of great importance for managing food production systems and to study their inter-relationships with water, geo-political, socio-economic, health, environmental, and ecological issues (Thenkabail et al., 2010). Furthermore, accurate development of all higher-level cropland products such as crop watering method (irrigated or rainfed), cropping intensities (e.g., single, double, or continuous cropping), crop type mapping, cropland fallow, as well as assessment of cropland productivity (i.e., productivity per unit of land), and crop water productivity (i.e., productivity per unit of water) are all highly dependent on availability of precise and accurate cropland extent maps. Uncertainties associated with cropland extent data affects the quality of all higher-level cropland products reliant on an accurate base map. However, precise and accurate cropland extent data are currently non-existent at the continental scale at a high spatial resolution (30-m or better). This lack of crop extent data is particularly true for complex, small-holder dominant agricultural systems of Africa. By mapping croplands at a high-resolution at the continental scale, the GFSAD30 project has resolved many of the shortcomings and uncertainties of other cropland mapping efforts.

The two most common methods for land-cover mapping over large areas using remote-sensing images are manual classification based on visual interpretation and digital per-pixel classification. The former approach delivers products of high quality, such as the European CORINE Land Cover maps (Büttner, 2014). Although the human capacity for interpreting images is remarkable, visual interpretation is subjective (Lillesand et al., 2014), time-consuming, and expensive. Digital per-pixel classification has been applied for land-cover mapping since the advent of remote sensing and is still widely used in operational programs, such as the 2005 North American Land Cover Database at 250-m spatial resolution (Latifovic, 2010). Pixel-based classifications such as maximum likelihood classifier (MLC), neural network classification (NN), decision trees, Random Forests (RF),and Support Vector Machines are powerful, and fast classifiers that help differentiate distinct patterns of landscape.
Both supervised and unsupervised classification approaches are adopted in pixel-based classifiers. However, per-pixel classification includes several limitations. For example, the pixel's square shape is arbitrary in relation to patchy or continuous land features of interest, and there is significant spectral contamination among neighboring pixels. As a result, per-pixel classification often leads to noisy classification outputs – the well-known "salt-and-pepper" effect. There are other limitations of pixel-based: 1. they fail to fully capture the spatial information of high resolution imagery such as from Landsat 30-m imagery, and 2. they often, classify the same field (e.g., a corn field) into different classes as a result of within field variability. This may often result in a field with a single crop (e.g., corn) classified as different crops.

For the creation of the GFSAD30AFCE data product, we used two supervised pixel-based classifiers (Pelletier et al., 2016, Tian et al., 2016, Shi and Yang, 2015, Huang et al., 2010): 1. Random Forests (RF's), and 2. Support Vector Machines (SVM's). Both of these have been widely used in agricultural cropland studies over the years

(Myint et al., 2011) and are considered powerful and ideal machine learning algorithms (Tian et al., 2016, Shi and Yang, 2015, Huang et al., 2010). A description of how to classify cropland extent of the African continent is provided in section 2.3 and its sub-sections.

In contrast, object-based classification involves identification of image objects, or segments, that are spatially contiguous pixels of similar texture, color, and tone (Congalton and Green, 2009). Whereas within field variability will be lost in object-based classification, the objects maintain the integrity of the farm field boundaries by grouping contiguous pixels, and overcoming the "salt-and-pepper" noise inherent in pixel-based classifications. We used recursive hierarchical segmentation (RHSeg) developed by Tilton et al., (2014, 2012) in this study. Section 2.3 and its sub-sections describe object-based RHSeg in detail. To take full advantage of these two approaches, the integration of pixel-based and the object-based analysis for large-area land cover mapping has been explored by several studies (Costa et al., 2014; Dingle Robertson and King, 2011; Malinverni et al., 2011; Myint et al., 2011) for limited study area regions. How these experimental methods perform in large areas is less well studied. Chen et al. (2015) presents an operational pixel-object-knowledge-based classification approach for producing 30-m Global Land Cover product (GlobeLand30). Even though GlobeLand30 was not designed to focus on cropland areas, it can readily be observed that there is high spectral heterogeneity within each land cover class and significant spectral confusion among different classes such as shrub and grass. In such cases, object-based segmentation has the potential to significantly improve the pixel-based classification results. As a result, we adopted two- pixel based supervised classifiers (RF and SVM) along with one object-oriented classifier (RHSEG) in producing GFSAD30AFCE. This fusion of methods would provide precise agriculture field boundaries under 30-m resolution (see overview of the methodology in Figure 1).

IV. Algorithm Description

An overview of the algorithm description provided in Figure 1. The methodology used in this project (Figure 1) is briefly described in this paragraph to provide an overview and presented in detail in subsequent sections of this ATBD document. The process (Figure 1) involved combining Sentinel-2 and Landsat-8 data along with SRTM 30-m data (see Table 2 below). The process included several well-designed steps (Figure 1). First, the data was pre-processed by cloud mask and gap-filling on Google Earth Engine (GEE). Second, seasonal moasics were created for two seasons in the areas where the majority of African crops are grown: Season 1 (January-June) and Season 2: July-December. Such a seasonal mosaic aided in achieving cloud free clear images of the continent. Each seasonal mosaic contained 11 bands as listed in Figure 1. Third, reference data were generated throughout Africa to train the RF and SVM algorithms. There are total of 8351 reference samples for this purpose. RHSeg was the applied by dividing Africa into a total of 1919 1 degree by 1 degree blocks (to aid in faster processing). Fourth, the results of the pixel-based RF and SVM algorithms were merged with the object-based RHSeg results to obtain the composite cropland extent product for Africa. Fifth, the composite cropland product of Africa was evaluated for accuracy using 1440 test samples. The process was iterated until adequate accuracies were attained. In this process the validation data was only available to the accuracy assessment team and was hidden from the production team. Finally, the GFSAD30AFCE product was made available on croplands.org as well as through LP DAAC via the LP DAAC Data Pool and NASA's Earthdata Search.

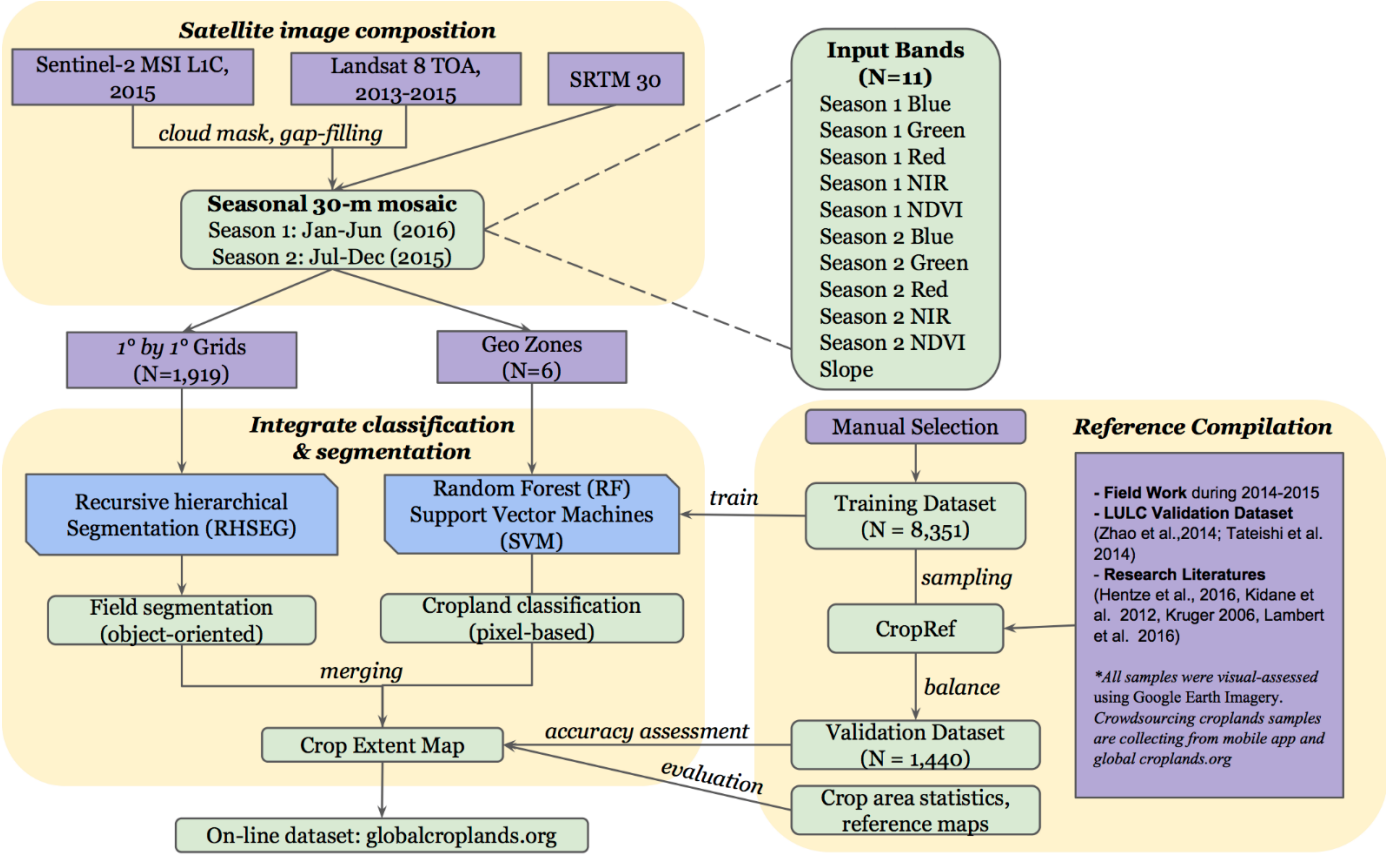

Figure 1. Flowchart of mapping methods for Sentinel-2 and Landsat-8 derived cropland Extent Product of Africa for the nominal year 2015.

a. Input data

i. Region Definition

Put a description of how Africa was defined – give a citation to the source from which the vector boundary was obtained. The study was conducted for the African continent that consists of 55 Countries (see Table 3, Figure 2, and Figure 7) that are recognized by African Union and\or United Nations. The country boundaries were determined by the Global Administrative Unit layers (GAUL) of United Nations (http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691&currTab=simple).

ii. Reference Samples

Reference data are required for both training and testing the machine learning algorithms (see section 2.3) as well as for validating the final products. Over 120,000 reference samples, spread across the world, were collected for this project and can be found at the following web site: https://croplands.org/app/data/search . Of these, there were over 9000 samples for Africa.

Reference training/testing data were obtained in the four ways. First, we gathered random samples by interpreting sub-meter to 5-meter very high spatial resolution imagery (VHRI) throughout Africa available to us from the National Geospatial Agency (NGA). There were a total 9,791 samples from VHRI spread across Africa. Second, reference samples were collected through several field campaigns by the project team from 2014-2015. These total 8,351 sample units (Figure 1) that capture crop properties including cropland location, irrigated *versus* rain-fed, crop intensity (e.g., single, double, continuous), and crop types. Third, some other global/region projects (Tateishi et al., 2014; Zhao et al., 2014) shared their valuable reference datasets. To incorporate these reference

data in our project, we converted their labeling system ("cross-walked") to be consistent with the labeling scheme of our project (Teluguntla, 2015). There were 651 samples spread across Africa from these sources. Fourth, reference samples were selected from a series of published literature sources for selected areas of Africa based on detailed studies using VHRI or high-resolution imagery such as Landsat (Haack et al., 2014; Kidane et al., 2012; Rembold et al., 2000; Shalaby and Tateishi, 2007; Were et al., 2013; Zucca et al., 2015). These studies provided an additional 500 samples that were uploaded to our reference database in globalcroplands.org.

iii. Satellite Imagery: Sentinel-2 and Landsat-8

In order to incorporate crop dynamics across the entire continent, Africa was divided (Figures 2 and 3) based on expert knowledge and literature review of the cropland calendar and precipitation patterns throughout the continent's agricultural systems (Hentze et al., 2016; Kidane et al., 2012; Kruger, 2006; Lambert et al., 2016; Motha et al., 1980; Waldner et al., 2016). Sentinel-2 five band multispectral imagery (MSI) 10-m and 20-m data (Drusch et al., 2012) was selected as the primary data source for the two main growing seasons (March – June; July - October) of Africa using data for the 2015-2016 growing seasons. Over 36,000 Sentinel-2 images were available for the African continent for this period and the cloud-free images were mosaicked using the median-value in the current growing season to ensure near-complete coverage. Even though Sentinel-2 results provided good coverage for Africa, some data gaps remained because of cloud and other data issues (Hollstein et al., 2016). Landsat-8 multispectral images at the resolution of 30 meters (Irons et al., 2012; Roy et al., 2014) were used as supplementary data for these data gaps, aiming to provide seamless 30-meter data for the two seasons for the entire study area. Data gaps were filled using a median data smoothing algorithm where missing data during a date was filled using data from previous and following dates of the imagery (Wang et al., 2017). Eventually, nominal 30-m imagery was generated for the entire study area (African continent).

The direct gap-filling of Sentinel-2 data with Landsat-8 data poses some technical challenges. The platform and sensor combinations differ in their orbital, spatial, and spectral configuration. As a consequence, measured physical values and radiometric attributes of the imagery were affected. For example, a root-mean-square error (RMSE) greater than 8% in the red band was found when comparing MSI and Landsat-7 simulated data, due to the discrepancies in the relative spectral response functions (RSRF) (D'Odorico et al., 2013). Werff and van der Meer (2016) compared Sentinel-2A MSI and Landsat 8 OLI Data, finding the correlation of their top-of-atmosphere (TOA) reflectance products was higher than their bottom-of-atmosphere (BOA) reflectance products. In addition, the combined use of the multi-temporal images requires an accurate geometric registration, i.e. pixel-to-pixel correspondence for terrain-corrected products. Both systems are designed to register Level 1 products to a reference image framework. However, the Landsat-8 framework, based upon the Global Land Survey images, contains residual geolocation errors leading to sensor-to-sensor misregistration of 38-m (Storey et al., 2016). Although both sensor geolocation systems use parametric approaches, whereby information concerning the sensing geometry is modeled and the sensor exterior orientation parameters (attitude and position) are measured, each uses a different ground control and digital elevation model to refine the geolocation (Languille et al., 2015; Storey et al., 2016). These misalignments vary geographically but not by more than $\pm$30-m.

Quantized and calibrated scaled Digital Numbers (DNs) for 4 MSI and OLI bands delivered as 16-bit unsigned integers were converted into Top-Of-Atmosphere (TOA) spectral reflectance (Gueymard, 2001).

$$T\,OA = \pi * L * d2/Esun * \cos(\theta) \qquad (1)$$

Where L is at-satellite radiance in W m-2 sr-1 mm-1, d is Earth-Sun distance in astronomical units, Esun is solar irradiance in W m-2 sr-1 mm-1, and θ is zenith angle in radians. Four MSI bands (blue, green, red, near-infrared) for every season and one slope band were extracted. The clouds were removed by using separate Quality Assessment (QA) band information available in the Sentinel-2 data. Landsat multi-bands were used only when Sentinel-2 data was missing because of cloud and the export spatial resolution is set to 30-m. We assessed that the mismatch between the geo-referencing of Landsat and Sentinel will always be within one pixel based on locating similar features from multiple locations across the continent in the Sentinel-2 and Landsat-8. Sentinel-2 has (B8(10-m) and B8A(20-m) bands for NIR range, where B8 is consistently lower than B8A due to different gain settings on the 10-m (B8) and 20-m (B8A). In order to match Landsat values better, the B8A band was used for NIR

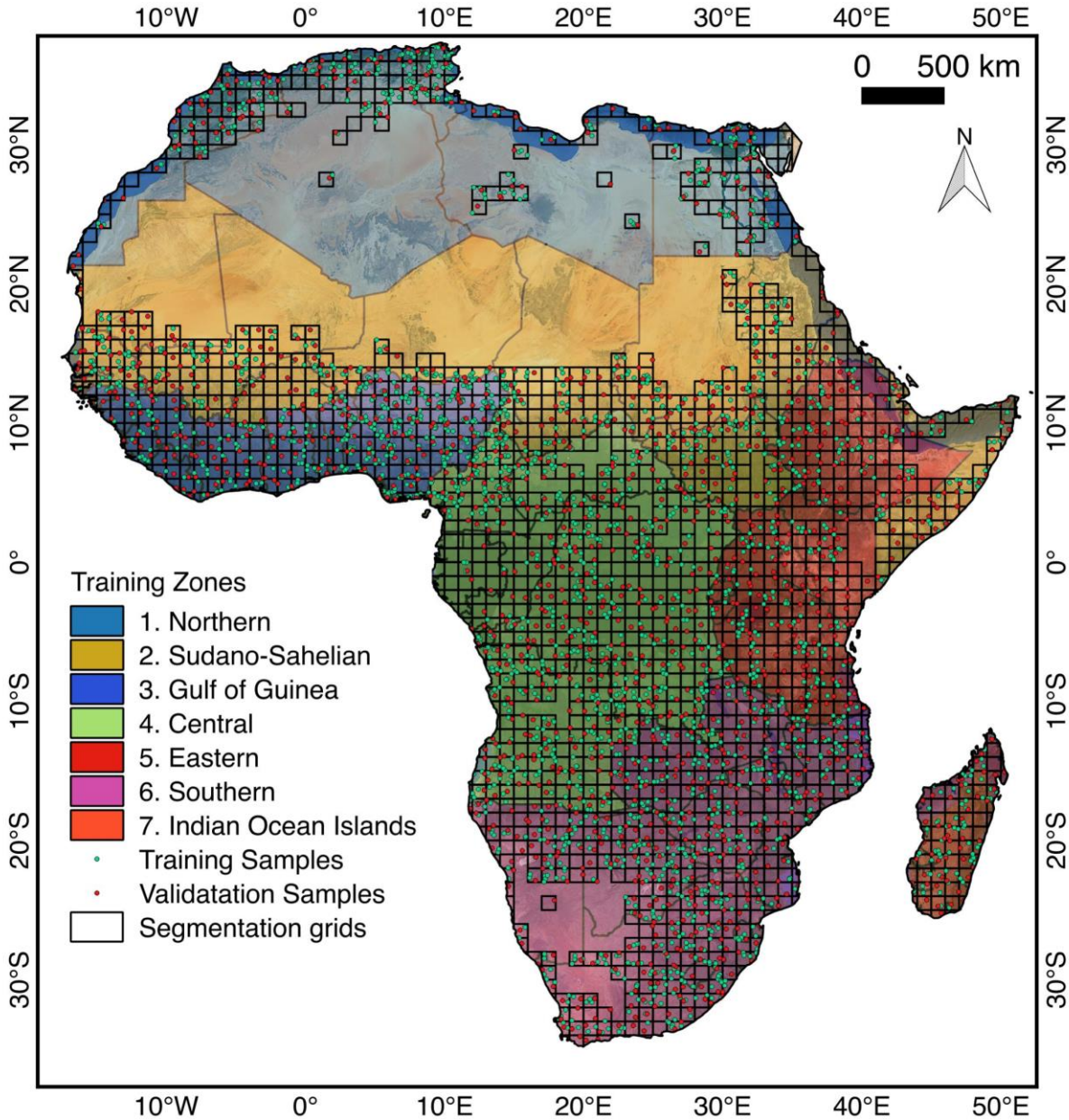

Figure 2. Stratification of the African continent into seven distinct refined FAO agro-ecological broad zones. The figure also shows the distribution of the reference training and validation data used in the machine learning algorithms.

A 30-m data cube (Figure 3) for classification was created in the following way: continental wall-to-wall mosaics were developed for two seasons: (Season 1: January-June 2016, Season 2: July-December, 2015). Five bands, blue, green, red, NIR and NDVI, were composited from Sentinel-2 and Landsat-8. From two seasons, we have 10 bands of data that were composited over two years (2015-2016). Note that each of the five bands are composited over each of the two seasons over two years- for a total of 10 bands from two seasons. In addition, we also included topographic information as an input variable in addition to the spectral information. We derived a slope surface from Shuttle Radar Topography Mission (SRTM) (Farr, 2007) digital elevation data at one arc-sec (approximately 30-m) resolution. These 11 bands (Figure 3) were organized as a Google Earth Engine (GEE) Image Collection object, which provides a programmable way for us to access the full archive for the entire African continent and run classification algorithms such as the Random Forest (RF) and Support Vector Machines (SVMs) for deployment on the GEE cloud environment.

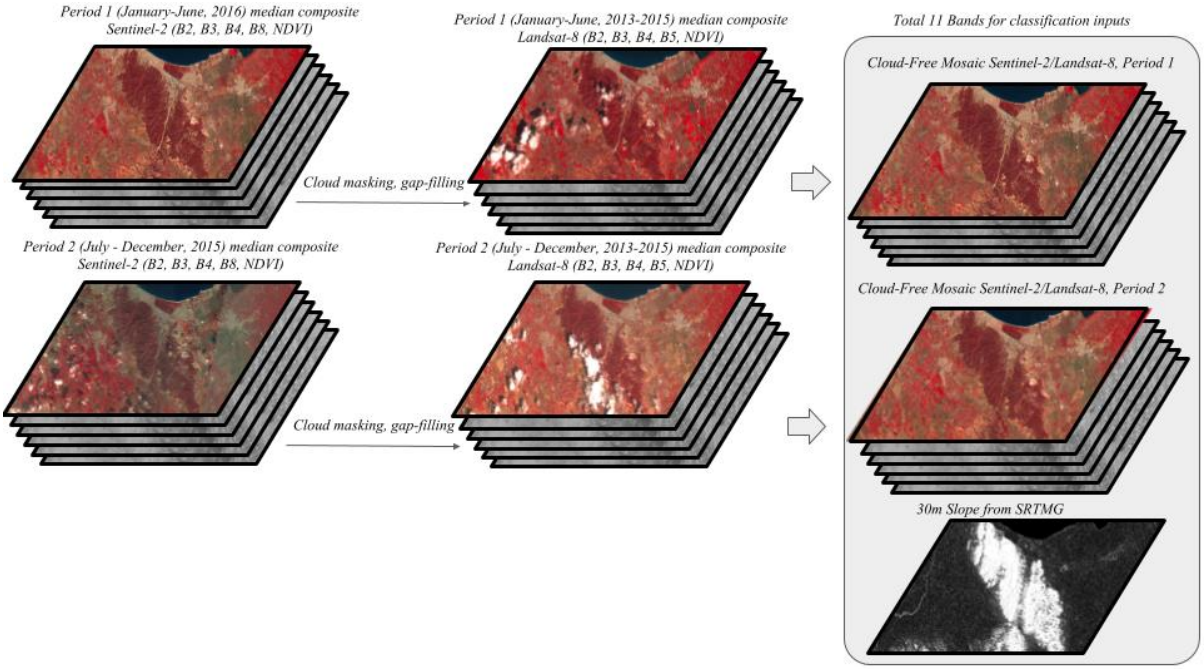

Figure 3. Data cube of 30-m for the entire African continent composited for two time-periods: (a) January-June, 2016, and (b) July-December, 2015 using time-series Sentinel-2 and Landsat-8 data. SRTM elevation data are also included. For each season five bands (B2, B3, B4, B8, NDVI of Sentinel-2 and Landsat-8, Table 2) are composited, taking the median value of a given pixel over the season.

Table 2. Characteristics of input Multi-Temporal Sentinel-2, Landsat data and slope band

Sensors	Period	Band	Use	Wavelength	Resolution	Provider
Sentinel-2 multi-spectral imaging (MSI) Level-1C, TOA reflectance	Season 1: January - June, 2016; Season 2: July - December, 2015	B2	Blue	490nm	10m	ESA
		B3	Green	560nm	10m	
		B4	Red	665nm	10m	
		B8A	Near Infrared	856nm	20m	
		NDVI			10m	
Landsat 8 top-of-atmosphere reflectance	Season 1: January - June, 2016; Season 2: July - December, 2015	B2	Blue	450 - 510nm	30m	USGS
		B3	Green	530 - 590nm	30m	
		B4	Red	640 - 670nm	30m	
		B5	Near Infrared	850 - 880nm	30m	
		NDVI			30m	
Shuttle Radar Topography Mission (SRTM) 30-m		Slope			30m	NASA/USGS

b. Theoretical description

1. Definition of Croplands

For all products within GFSAD30, cropland extent was defined as, "lands cultivated with plants harvested for food, feed, and fiber, including both seasonal crops (e.g., wheat, rice, corn, soybeans, cotton) and continuous plantations (e.g., coffee, tea, rubber, cocoa, oil palms). Cropland fallow are lands uncultivated during a season or a year but are farmlands and are equipped for cultivation, including plantations (e.g., orchards, vineyards, coffee, tea, and rubber" (Teluguntla et al., 2015). Cropland extent includes all planted crops and fallow lands. Non-croplands include all other land cover classes other than croplands and cropland fallow (Figure 4).

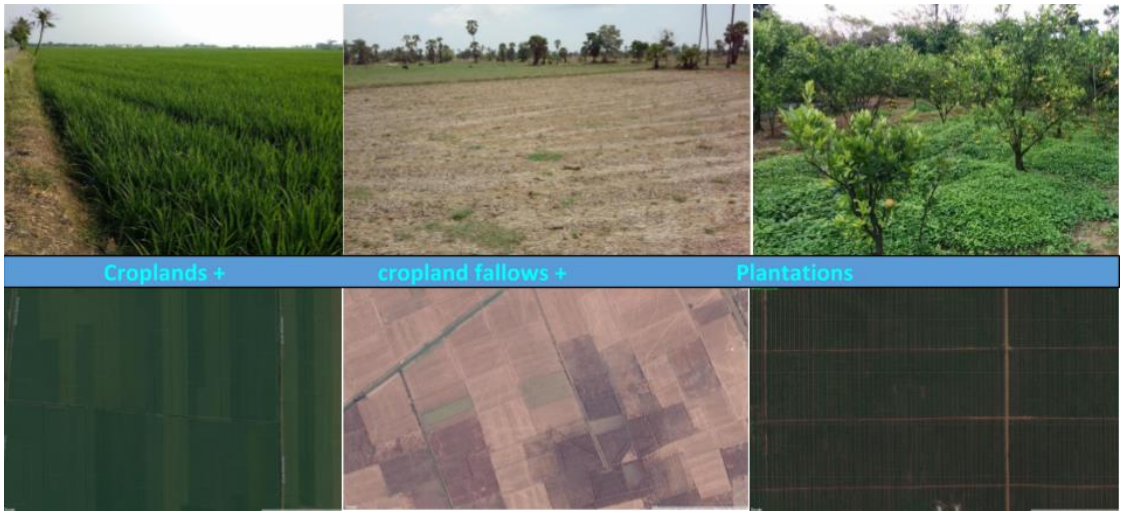

Figure 4. Illustration of definition of cropland mapping. Croplands included: (a) standing crop, (b) cropland fallows, and (c) permanent plantation crops.

2. Algorithms

The study used three machine learning algorithms to create the cropland extent product. Two pixel-based supervised classifiers, Random Forest (RF) and Support Vector Machines (SVMs), and one object-oriented classifier, Recursive Hierarchical Image Segmentation (RHSEG). These algorithms are described in detail below. Africa was stratified into seven separate refined FAO agro-ecological zones (Northern, Sudano-Sahelian, Gulf of Guinea, Central, Eastern, Southern and Indian Ocean Islands; Figure 2) to facilitate optimal classification.

c. Practical description

1. Random Forest (RF) Algorithm

The Random Forest classifier is more robust, relatively faster in speed of classification, and easier to implement than many other classifiers (Pelletier et al., 2016). The Random Forests classifier uses bootstrap aggregating (bagging) to form an ensemble of decision trees (Pelletier et al., 2016) by searching random subspaces from the given data (features) and the best splitting of the nodes by minimizing the correlation between the trees

All supervised pixel-based classifications are heavily dependent on the input training samples selected. In order to discriminate croplands under various environments and conditions, the sample size of the initial training dataset needs to be large, especially in complex regions. All samples were selected to represent a 90-m x 90-m polygon. First, extensive field campaigns were conducted in Africa during the 2014-2015 crop growing season to collect data on precise cropland locations as well as non-cropland locations. This effort led to collection of 1,381 samples spread across Africa. Second, we compiled the ground data from other reliable sources. Third, sub-meter to 5-m very high spatial imagery was used to generate croplands *versus* non-cropland samples using multiple interpreters across Africa. Approximately 7,000 data samples were used generated from these interpretations. An iterative sample selection procedure was implemented with the following steps for training the Random Forest (RF) machine learning algorithm is described below (also see logical flow in Figure 1):

1. Build the Random Forest classifier on Google Earth Engine (GEE) cloud computing environment using existing training samples for each of the seven zones (e.g., Figure 2). Initially we began with a small number of samples and slowly increased the sample size until we reached a high degree of accuracy and the accuracy plateaued at certain sample size. Our experiments showed that for each zone, this plateau was reached at approximately 250 samples beyond which the increase in sample size resulted in insignificant increases in the accuracies of classification results;

2. Classify the 30-m seasonal mosaics (Figure 3) for each of the seven zones (Figure 2) using the Random Forest algorithm in GEE cloud;
3. Visual assessment of classification results were compared with existing reference maps as well as sub-meter to 5-m very high spatial resolution imagery (VHRI). The process (Figure 1) was iterated until sufficient correspondence was achieved;
4. Added (or balancing, see Figure 1) 'crop' and 'non-crop' samples in areas that were not covered using reference data obtained from the sub-meter to 5-m very high spatial imagery from Google Earth Imagery. For locations where interpretations were challenging (fallow-land or abandoned fields), historical Landsat 5/7/8 Images and ground data were also used.
5. Loop step 1-4 by progressively increasing the training dataset until classification becomes stable.

The number of iterations required for the training sample selection is a function of the complexity of the area. In the rainfed areas of central Africa like Tanzania, the rainfed cropland is highly mixed with natural vegetation and barren land. In this situation, the iterative selection was looped 4~5 times to improve the classification results.

2. Support Vector Machines (SVMs) Algorithm

Support Vector Machines (SVMs) are a pattern recognition method that can address problems such as small sample size, nonlinearity, high dimensionality, and local minima that locates within a set of points which may or may not be a global minimum and it is not the lowest value in the entire set (Vapnik and Vapnik, 1998). SVMs have been widely used in remote sensing studies and can improve classification accuracy as compared to traditional classification methods, such as the maximum likelihood classification (Foody and Mathur, 2006; Pal, 2007). SVMs are generally considered to be superior to Random Forest for a number of reasons including: (1) their ability to gather data from kernels for a more nuanced assessment (e.g., linear, polynomial) of the map classes and (2) the ability to apply the knowledge generated by SVM hyperplanes generated from small, intelligently selected training samples to the rest of the data (Shi and Yang, 2015). SVMs project raw input data into a higher dimensional space to increase the separability between different classes when they cannot be appropriately separated by a linear hyperplane (where data are transformed and projected in multi-dimensional feature space) that maximizes the margin between the two classes. This transformation is realized through different kernel functions and training samples, which cause more scatter after projection into a higher dimensional space.

3. Combining Random Forest & Support Vector Machines for optimal results

First, we run RF algorithm (section 2.3.1) to classify and separate croplands versus non-croplands. However, RF has over-fitting problem (Figure 5, middle image), meaning some of the non-croplands can get mapped as croplands in certain complex areas (e.g., mountainous terrain, fragmented landscape where croplands and non-croplands intermix in various ways). So, RF often has commission errors in these difficult landscapes. So, in such areas of significant commission errors, we mask out that area and run SVM. The SVM resolves the over-fitting problem by separating croplands from non-croplands (Figure 5, right image) much easily than RF.

SVM and RF are two widely used machine-learning algorithms that have proven to be capable of handling complex classifications with a large number of input features. SVM and RF were selected as two different strategies for the use of training samples (global and regional samples based on a spatial-temporal selection criterion) and were performed in this project (Yu et al., 2013). The combination of Random Forest (RF) and Support Vector Machine (SVM) classifiers has proven to: (i) enhance crop classification accuracy and (ii) provide spatial information on map uncertainty (Löw et al., 2012). In addition, experimental results indicate that this hybrid classifier improves classification accuracy in comparison to the single classifiers.

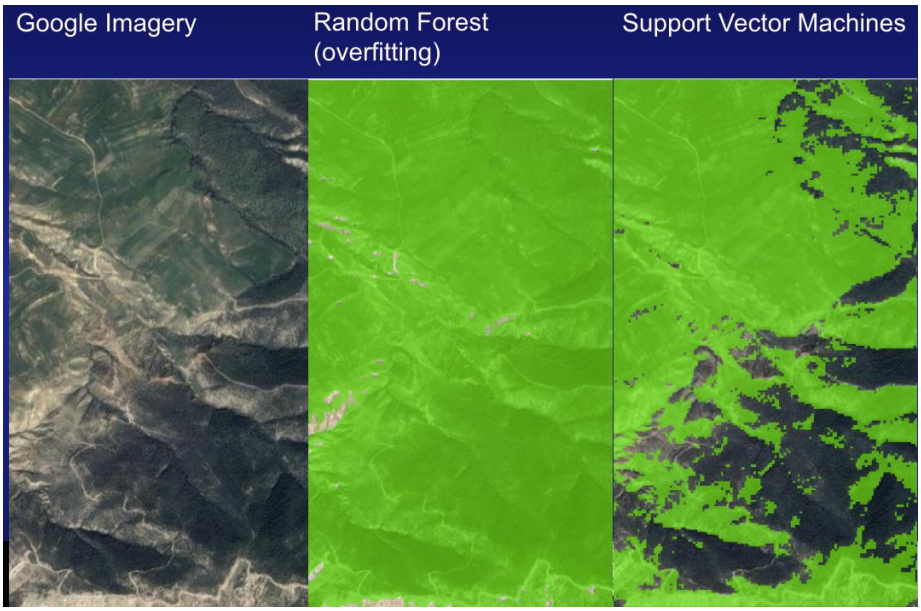

Figure 5. Illustration of overfitting (mapping some of the non-croplands also as croplands) by the Random Forest (RF) algorithm (center image) is overcame by Support Vector Machines (SVM's) (rightmost image). Compare these two results with the reference sub-meter to 5-m imagery (leftmost image). RF requires greater number of training samples, but is more robust. However, it has an over-fitting problem. The SVM requires less training data to classify, and avoids overfitting, but is more complicated than RF. Combining the two can help reduce uncertainties in cropland mapping.

4. Recursive Hierarchical Image Segmentation (RHSeg)

Image segmentation splits an image into separated regions or objects depending on parameters specified (Im et al., 2008; Stow et al., 2008). A group of pixels having similar spectral and spatial properties is considered an object in the object-based classification prototype.

The recursive hierarchical segmentation (RHSeg) segmentation algorithm (Tilton et al., 2012) was selected to segment the imagery. RHSeg is an approximation of the Hierarchical Segmentation (HSeg) algorithm that enables the processing of large images with HSeg. RHSeg recursively subdivides an image into smaller subsections that can be processed by HSeg in a reasonable amount of time. HSeg uses the best merge region growing combined with non-adjacent region merging to produce an image segmentation. HSeg provides a choice of region dissimilarity functions to use in defining these best merges. In this study, we selected dissimilarity option 6, which is a criterion based on mean squared error between the region mean value and the original image data (Tilton et al, 2012). HSeg also provides an option to use image-edge feature information to influence the merging of spatially adjacent regions to inhibit merging regions across a boundary with high edge feature values (Tilton and Pasolli, 2014). The relative importance of spatially adjacent region object merging versus non-adjacent region object merging is controlled by a "spectral clustering" weighting factor. We used a value of 0.5 for this parameter, which prioritized spatially adjacent region merging over non-adjacent merging by a factor of 2.

HSeg and RHSeg are designed to produce a hierarchical set of image segmentations that can be visually inspected or otherwise analyzed to determine the appropriate level of segmentation detail for a particular application. After visually inspecting the RHSeg hierarchical segmentation results from eleven representative images, we judged that the segmentation results at merge thresholds 7.5 and 15.0 produced appropriate levels of segmentation detail. (We made our selection in terms of merge thresholds instead of number of regions because we noted that some image scenes had a significant percentage of water pixels or were masked out due to clouds, and we realized that more consistent results would be obtained by selecting results from the RHSeg segmentation hierarchy based on

merging thresholds instead of the number of regions.) These RHSeg segmentation results provided detailed agriculture field borders at 30-m resolution that were readily integrated with the pixel-based classification results.

5. Integration of pixel-based classification and hierarchical segmentation

Per-pixel classification has several limitations. For example, the pixel's square shape is arbitrary in relation to patchy or continuous land features of interest. Also, there is a relevant exchange of spectral contamination among neighboring pixels (known as the modulation transfer function). As a result, pixel-based classification may generate a large number of misclassified pixels (i.e., the "salt-and-pepper" effect) due to the spectral confusion between land cover types and spectral diversity within the same land cover type. Alternatively, object-based segmentation can preserve membership information about whether pixels come from the same field. To integrate the pixel-based classification results with the segmented objects, we modified a per-pixel classification (e.g., from a Random Forest classifier) in the following manner (Figure 6): For each region object, if 85% or more of the pixels in the region were classified as crop, we labeled all of the pixels in the region as crop. If 15% or less of the pixels in the region were classified as crop, we labeled all of the pixels in the region as non-crop. We left unchanged the pixel classification of the pixels in the remaining regions.

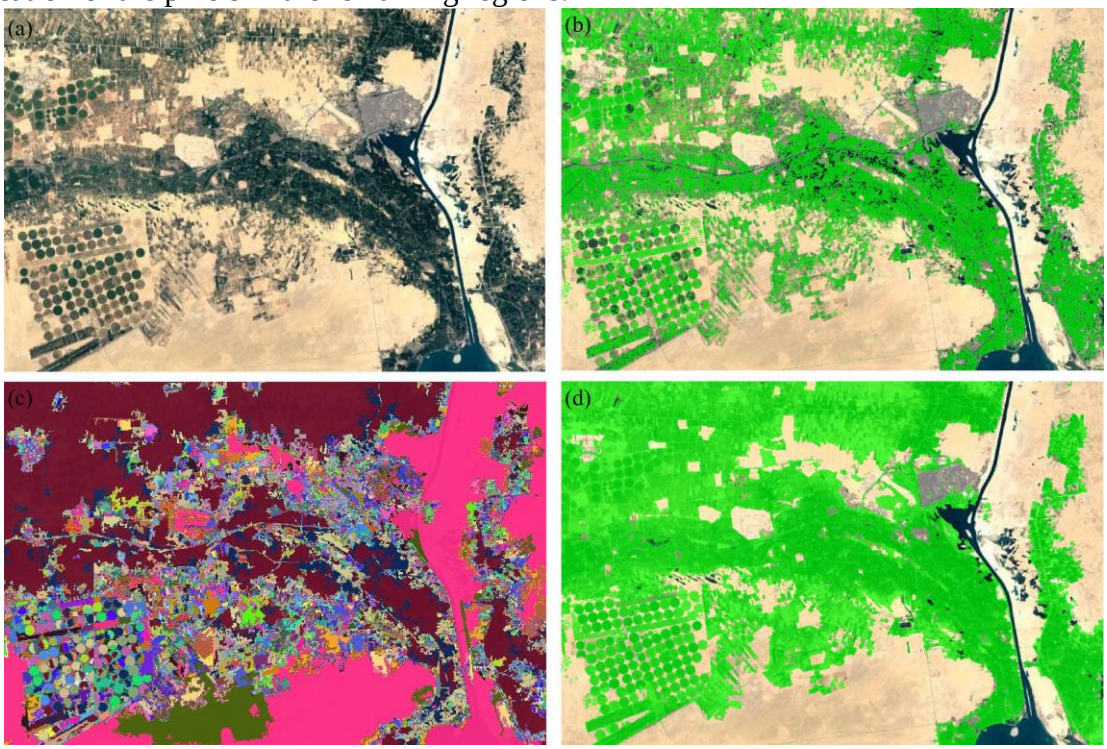

Figure 6. The example of (a) a true color sub-meter to 5-m Google Earth Imagery is used for reference and compared with: (b) combined results from the pixel-based Random Forest (RF) and support vector machine (SVM) classifier, (c) the object-based RHSeg image segmentation result with edge-based processing window artifact elimination, and (d) the merged results with RHSeg segmentation result with pixel-based Random Forest and support vector machine classification, which is spatially more consistent.

The merging of spatially non-adjacent regions in RHSEG leads to heavy computational demands. In order to expand its capability from a regional-size (maximum 8,000 x 8,000 images) to continental scale, a gridding scheme was applied to subset the study area into much smaller pieces. Each scene was about 4,000 columns by 4,000 rows in size. We applied the above-proposed method to continental Africa. The Random Forest classification was carried out in seven geographical zones (Figure 2) separately. There were 1919 30-m resolution data sets covering the non-dessert areas of Africa using a 1° by 1° latitude/longitude grid produced (Figure 2). These data sets were two period mosaics (period 1: Jan-Jun 2016, period 2: Jul-Dec 2015) with NIR, red, green and blue spectral bands from each season. The source image data was Landsat 8 TOA Reflectance fused with Sentinel-2

MSI, Level-1C. Creating these data sets took about 74 hours of wall clock time using 64 CPUs and the total size of the datasets was about 800 gigabytes.

6. Programming and codes

The two pixel-based supervised machine learning algorithms (RF and SVM's) were coded on Google Earth Engine (GEE) using Python and Java Scripts using Application Programming Interface (API). The RHSeg is coded and computed on NASA supercomputers, including Goddard's Discover machine and Ames' Pleiades machine. The codes are made available in a zip file and are available for download along with this ATBD.

7. Results

The machine learning algorithms (RF, SVM, RHSeg), discussed in previous sections, were trained to separate croplands versus non-croplands for each of the 6 zones (Figure 2) based on the previously described reference data. The three machine learning algorithms were then run on the Google Earth Engine (GEE) cloud computing environment using the Sentinel-2 and Landsat-8 collection for each of the 7 zones to separate croplands versus non-croplands. The process was iterated and knowledge in the algorithms tweaked several times, before producing the final, accurate results of croplands versus non-croplands. This process resulted in the global food security-support analysis data @ 30-m cropland extent for Africa (GFSAD30AFCE) product (Figure 7). This product is publically available through the Land Processes Distributed Active Archive Center (LP DAAC). The same dataset is also available for visualization at https://croplands.org/app/map. Full resolution of 30-m cropland extent can be visualized in croplands.org by zooming-in to specific areas as illustrated in panels (b) and (c) of Figure 7. For any area in Africa, croplands can be visualized by zooming into specific areas in croplands.org. The background sub-meter to 5-m imagery, available for the continent on the Google Earth, helps evaluate the quality of the cropland extent product ("zoom in" and "toggle" cropland "on" and "off" to see the sub-meter to 5-m imagery in the background.

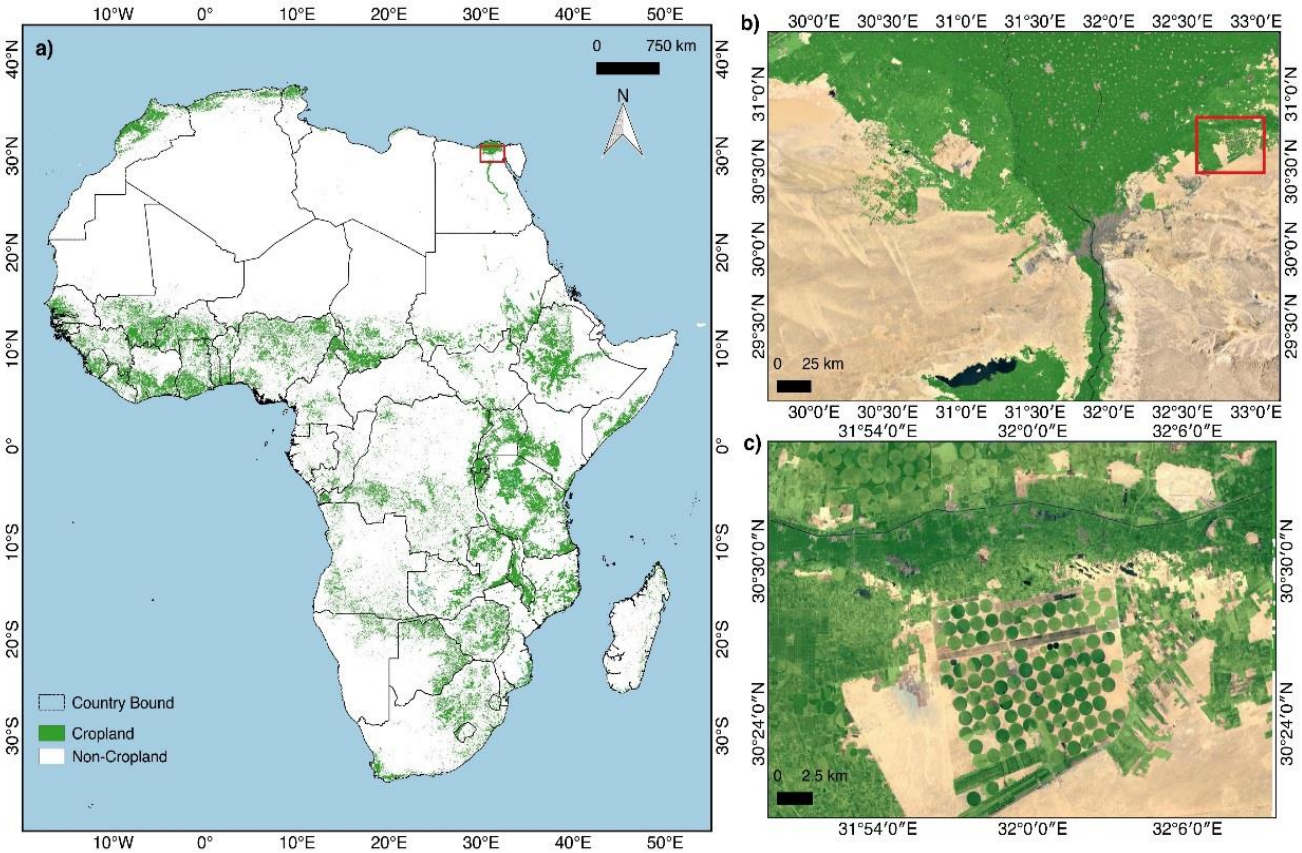

Figure 7. Cropland Extent Product at 30-m for the African continent (GFSAD30AFCE). This product is made available for visualization @: croplands.org.

8. Cropland areas of Africa

Table 3 shows the country-by-country cropland area statistics for all 55 African Countries generated in this study and compares them with several other sources such as the national census data base MIRCA2000 (Stefan Siebert and Portmnn, personal communication; Portmann, 2010) which was also updated in the year 2015, The Food and Agricultural Organization (FAO) of United Nation's compiled statistics, the MODIS derived GFSAD250 (Xiong et al., 2017), the MODIS 500-m derived cropland areas from GRIPC (Salmon et al., 2015), and the GLC30 (Chen et al., 2015) . Overall, as derived from GFSAD30AFCE estimates, the entire African continent had total net cropland area (TNCA) of 313 Mha. Our recent MODIS-based GFSAD250 estimated TNCA as 296 Mha. All other studies estimated TNCA between 211-232 Mha with MIRCA estimates for nominal year 2015 estimating 232 Mha. The overwhelming proportion (94%) of the cropland areas (Table 3) were in just 27 of the 55 countries. Compared to MIRCA2000, the GFSAD30AFCE (Figure 8) sometimes overestimates and sometimes underestimates cropland areas.

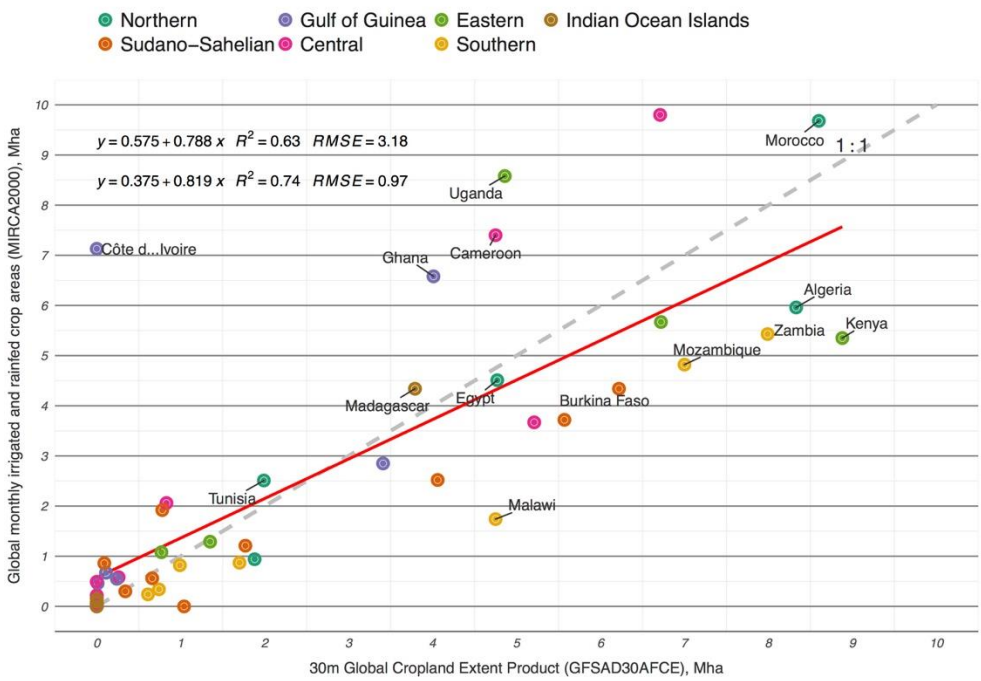

Figure 8. Country-by-country scatterplot of GFSAD30AFCE and MIRCA2000 (Portmann, 2010).

On average, the GFSAD30AFCE had much higher estimates -- higher by about 35-40% relative to statistical data from MIRCA and FAO (Table 3). There are several reasons for these significant differences in areas between GFSAD30AFCE when compared with MIRCA and FAO estimates. These include:

(i) The ability of the higher spatial resolution 30-m derived croplands GFSAD30AFCE map to capture fragmented croplands;

(ii) The ability of the 30-m derived croplands GFSAD30AFCE map to account for actual areas when compared with sub-pixel areas of lower resolution imagery derived cropland products (e.g., GRIPC, GLC, GFSAD250);

(iii) Differences in how cropland data are gathered\estimated\calculated. The 30-m derived cropland GFSAD30AFCE map provides objective estimates relative to how other statistical data were obtained. Statistical data of countries are reported by countries based on a wide range of methods, techniques, and data used. For example, FAO compiles the statistics reported by individual countries, which are based on national censuses, agricultural samples, questionnaire-based surveys with major agricultural

producers, and independent evaluations (FAO, 2006 and The World Bank, 2010). Since each country has its own data collection mechanism, differences in data gathering, resource limitations, and lack of objectivity in many countries (due to resource limitations) results in data quality issues, particularly in Africa. For example, in 2008/09 in Malawi, cropland extent was estimated by combining household surveys with field measurements derived from a "pacing method" in which the size of crop fields was determined by the number of steps required to walk around them (Dorward, 2010). Some countries (e.g., USA, Canada) use remote sensing, others use subjective eye estimates, and many others many not have resources to maintain proper statistics and report a reliable number;

(iv) Definition issues. Not every country adheres to same definitions of croplands. Our study used the TNCA definition to include planted crops along with croplands left fallow as well as permanent crops such as plantations (e.g., olive, fruit trees, vineyards, coffee and tea plantations, oil palm plantations etc). Many countries use similar definitions while others use different definitions (e.g., leaving out cropland fallow).

(v) Uncertainties inherent in all estimates. One can expect uncertainties in cropland areas maps (e,g., Figure 7) or areas estimated from different sources (e.g., Table 3) as a result of definitions, data used, methods adopted, and reporting mechanisms (e.g., FAO mostly reports official areas reported by the Countries). GFSAD30AFCE uncertainties are gauged by the error matrices (Table 4). In GFSAD30AF uncertainties in cropland estimates mainly arose from three sources: (a) aquaculture, (b) greenhouses, and (c) managed farmlands. Aquaculture are part of the farming system, especially adjoining rice fields in delta's in SE Asia and South Asia and were not present in Africa and hence is a non-issue in GFSAD30AFCE product. However, some uncertainties are expected in this as some of the water bodies adjoining rice fields can be confused as aquaculture. Green houses are croplands, but difficult to map. We used sub-meter to 5-m very high-resolution imagery in areas dominated by greenhouses to include these in croplands and were successful. However, some uncertainties arise when greenhouses are fragmented and isolated in remote areas. A major difficulty is in separating croplands from managed pasture.

There were 7 "outlier" countries in cropland area estimation when the GFSAD30AFCE product was compared to MIRCA2000 (Figure 8). There are many reasons for this. For example, GFSAD30AF methods and approaches were purely remote sensing based as opposed to predominantly survey-based statistics used in MIRCA2000. MIRCA2000 is a derived gridded dataset based on the FAOSTAT database (Portmann, 2010).

Table 3. Net cropland areas (NCAs) of Africa based on 30-m cropland product and comparison with other cropland products.

Country	Land (Mha)	Total Net Cropland Area (TNCA, Mha)					
	FAO-GAUL	MIRCA2000	This Study: GFSAD30AFCE	FAO Cultivated area (2002)	GRIPC	GFSAD250	GLC30
Resolution	--	variable	30-m	variable	500-m	250-m	30-m
Nigeria	90.56	38.62	35.67	33.00	39.30	14.05	28.20
Ethiopia	112.76	11.09	25.70	10.67	14.23	19.63	21.72
Sudan	186.88	18.40	22.74	16.65	10.43	9.09	19.94
Tanzania	93.98	5.67	22.57	5.10	4.57	28.81	18.25
South Africa	122.00	15.70	19.91	15.71	10.99	13.06	15.20
Congo DRC	232.94	9.80	16.32	7.80	11.41	22.87	4.47
Mali	125.26	4.84	12.78	4.70	9.81	10.66	5.16
Zimbabwe	39.07	3.53	12.31	3.35	0.10	10.75	8.97
Zambia	75.12	5.43	9.70	5.29	0.23	15.31	6.49
Kenya	59.34	5.35	9.23	5.16	5.74	8.91	8.38
Morocco	67.77	9.68	8.98	9.28	6.18	5.76	8.06
Algeria	231.27	5.96	8.81	8.27	3.99	3.78	7.73
Niger	118.12	14.53	8.45	4.50	1.06	0.24	6.60
Mozambique	78.57	4.82	8.42	4.44	1.30	11.76	5.80
Côte d'Ivoire	32.07	7.13	7.86	6.90	9.53	9.02	1.44
Burkina Faso	27.39	4.34	7.37	4.40	10.39	7.95	4.19
Uganda	24.13	8.58	7.19	7.20	9.37	7.69	6.24
Angola	124.71	3.67	6.32	3.30	2.17	7.94	4.29
Chad	127.09	3.72	6.31	3.63	7.13	11.07	4.80
Cameroon	46.50	7.40	5.20	7.16	4.45	5.37	1.45
Malawi	11.85	1.74	5.20	2.44	0.46	5.83	3.76
Tunisia	15.50	2.51	5.05	4.91	2.21	1.66	4.53
Egypt	98.22	4.51	4.99	3.42	3.21	4.11	4.31
Madagascar	58.98	4.34	4.63	3.55	4.02	2.55	2.07
Ghana	23.86	6.58	4.62	6.33	6.72	8.73	2.16
Senegal	19.52	2.52	4.33	2.51	5.65	2.59	3.32
Benin	11.52	2.85	3.81	2.82	2.44	2.97	2.95
Togo	5.67	2.63	2.24	2.63	1.53	1.65	1.73
Libya	161.52	0.94	2.08	2.15	0.64	0.45	1.68
Somalia	63.26	1.21	2.05	1.07	1.95	2.52	1.54
Botswana	57.84	0.87	1.90	0.38	0.02	9.75	0.84
Rwanda	2.53	1.29	1.42	1.39	1.68	1.63	1.20
South Sudan	62.43	0.00	1.26	0.00	0.00	4.13	0.00
Namibia	82.39	0.82	1.23	0.82	0.00	3.80	0.80
Central African Republic	62.03	2.06	1.02	2.02	0.61	1.18	0.69
Guinea	24.47	1.92	0.97	1.54	0.85	5.32	0.64
Lesotho	3.05	0.34	0.85	0.33	0.08	0.54	0.62
Burundi	2.72	1.08	0.85	1.35	1.26	1.80	0.67
Eritrea	12.25	0.56	0.73	0.50	0.16	0.33	0.56
Swaziland	1.74	0.24	0.69	0.19	0.19	0.38	0.52
Gambia	1.06	0.30	0.38	0.26	0.63	0.57	0.26
Congo	34.22	0.58	0.31	0.24	3.40	2.35	0.22
Guinea-Bissau	3.36	0.55	0.31	0.55	0.14	0.76	0.20
Sierra Leone	7.23	0.67	0.13	0.60	0.87	2.73	0.09
Mauritania	103.89	0.86	0.11	0.50	0.03	0.07	0.08
Liberia	9.58	0.46	0.01	0.60	0.21	2.44	0.01
Cape Verde	0.41	0.06	0.00	0.05	0.02	0.00	0.00
Gabon	26.15	0.49	0.00	0.50	0.94	1.07	0.00
Sao Tome and Principe	0.10	0.03	0.00	0.05	0.00	0.01	0.00
Equatorial Guinea	2.69	0.22	0.00	0.23	0.01	0.13	0.00
Comoros	0.17	0.08	0.00	0.13	0.04	0.00	0.00
Djibouti	2.17	0.00	0.00	0.00	0.00	0.00	0.00
Mauritius	0.20	0.16	0.00	0.11	0.12	0.00	0.05
Saint Helena	0.04	0.00	0.00	0.00	0.00	0.00	0.00
Seychelles	0.05	0.00	0.00	0.01	0.00	0.00	0.00
Total	**2988**	**232**	**313**	**211**	**202**	**296**	**223**

V. Calibration Needs/Validation Activities

For this assessment, 1,750 reference samples that were collected independently of any reference training and testing samples used by the mapping team were used. Error matrices were generated for each of the seven zones separately and for the entire African continent providing producer's, user's, and overall accuracies (Story and Congalton, 1986, Congalton, 1991, and Congalton and Green, 2009).

Table 4. Independent Accuracy Assessment of 30-m Cropland Extent Map for Africa. Accuracies were assessed for each of the 7 zones as well as for the entire continent.

Zone 1, % of TNCA* = 9.1%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	43	5	48	89.6%
	No-Crop	4	198	202	98.0%
Total		47	203	250	
Producer Accuracy		91.5%	97.5%		
Overall Accuracy		96.4%		Fscore	0.91

Zone 2, % of TNCA* = 26.4%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	21	8	29	72.4%
	No-Crop	8	213	221	96.4%
Total		29	221	250	
Producer Accuracy		72.4%	96.4%		
Overall Accuracy		93.6%		Fscore	0.72

Zone 3, % of TNCA* = 21.7%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	37	21	58	63.8%
	No-Crop	2	190	192	99.0%
Total		39	211	250	
Producer Accuracy		94.9%	90.0%		
Overall Accuracy		90.8%		Fscore	0.76

Zone 4, % of TNCA* = 6.2%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	8	7	15	53.3%
	No-Crop	1	234	235	99.6%
Total		9	241	250	
Producer Accuracy		88.9%	97.1%		
Overall Accuracy		96.8%		Fscore	0.67

Zone 5, % of TNCA* = 16.6%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	44	17	61	72.1%
	No-Crop	5	188	193	97.4%
Total		49	205	254	
Producer Accuracy		89.8%	91.7%		
Overall Accuracy		91.3%		Fscore	0.80

Zone 6, % of TNCA* = 19.9%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	22	9	31	71.0%
	No-Crop	4	215	219	98.2%
Total		26	224	250	
Producer Accuracy		84.6%	96.0%		
Overall Accuracy		94.8%		Fscore	0.77

Zone 7, % of TNCA* = 0.1%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	17	7	24	70.8%
	No-Crop	11	215	226	95.1%
Total		28	222	250	
Producer Accuracy		60.7%	96.8%		
Overall Accuracy		92.8%		Fscore	0.65

All Zones, % of TNCA*=100%		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	176	81	257	68.5%
	No-Crop	29	1464	1493	98.1%
Total		205	1545	1750	
Producer Accuracy		85.9%	94.8%		
Overall Accuracy		93.7%		Fscore	0.76
Weighted Accuracy		94.5%			

Note: * TCA (Total Croplands Area) = 313 Mha

** The all-zones Weighted Accuracy is weighted by proportion of croplands in each zone

For the entire continent, the weighted overall accuracy was 94.5% with producer's accuracy of 85.9% and user's accuracy of 68.5% for the cropland class (Table 4). This means, for the cropland class of Africa, the errors of omissions were 14.1% and errors of commissions were 31.5%. This shows that we have missed only 14.1% of the croplands in the continent. When considering all 7 zones, the overall accuracies ranged between 91-96% (rounded off to nearest integer), producer's accuracies ranged between 61-97%, and user's accuracies ranged between 53-90% (Table 4). Zones that included a larger proportion of croplands had high overall, user's, and producer's accuracies. These results clearly imply the high level of confidence in differentiating croplands from non-croplands for the African continent.

The overall accuracies were highest in zone 1 (Table 4). Zone 1 is characterized by agricultural areas immediately adjacent to desert, with most of the croplands concentrated along the coast. High producer's accuracies across zones suggest than few croplands were omitted during the mapping process. On the other hand, high user's accuracies across zones suggest than croplands were rarely mapped (or committed) in error. The three machine learning algorithms (RF, SVMs, and RHSeg) were optimized to map the maximum extent of croplands. To some extent, this decision increases commission errors. In summary, the producer's accuracy in all but two zones were 85% or higher, which clearly indicates that croplands have been mapped with high accuracy (Table 4).

VI. Constraints and Limitations

GFSAD30AFCE product mapped the croplands of Africa @ nominal 30-m, which is the best-known resolution for cropland mapping over such large area as African continent covering all 55 Countries. It also has high levels of accuracies with weighted overall accuracies of 94.5%, Fscore of 0.76, Producer's accuracy of 85.9% and user's accuracy of 68.5%.

A producer's accuracy of 85.9% for the cropland class means an error of omission of 14.1%. This means 14.1% of the continental croplands were missing in the product. A user accuracy 68.5% for the cropland class for the continent means there is an error of commission of 32.5%. This means, 32.5% of non-croplands are mapped as croplands. We tweaked the machine learning algorithms (section IV) to maximize capturing as much croplands as feasible automatically. In this process, some non-croplands get mapped as croplands as well. This is a preferred solution in order not to miss croplands or only to miss them minimally. As a compromise mapping some non-croplands as croplands becomes unavoidable.

There are numerous issues that cause uncertainties and limitations in cropland extent product. Some of these issues are discussed here. First, temporal coverage. Five to ten-day Sentinel-2 and 16-day Landsat-8 coverage put together, there is substantial temporal coverage. Yet, if we look at Figure 3, we were only able to achieve seasonal cloud-free or near cloud-free mosaics of the entire African continent. This is not surprising given the such a large area involved and frequent cloud (e.g., frequent clouds over the Congo rainforests) or dust (e.g., Harmattan dust blown off Sahara) across the continent. As a result, if we were to have daily coverage over an area (e.g., like MODIS) then it becomes feasible to have more frequent (e.g., monthly or bi-monthly composites) temporal coverage of the continent that will help advance cropland mapping at improved accuracies. Second, there is a need for greater understanding of the Sentinel-2 and Landsat-8 data on how well they are correlated and in efforts to achieve better harmonization of data from two different sensors. Third, is the limitation of the reference training and validation data. In this project, we already have large training and validation data compared to any previous work as described in various previous sections. Nevertheless, much wider and extensive field visits to different parts of the continent will be helpful in better understanding of the issues involved and as a result better mapping. For example, slash and burn croplands in the rainforests or agroforest driven croplands in many parts of Africa, helps us in better define, understand, and map croplands. These and a better understanding of croplands through field visits as well as understanding of host of other issues (e.g., various types of irrigated and rainfed croplands, croplands in desert margins, various types and ages of cropland fallows) all will help improve cropland mapping. Greatest difficulties in cropland mapping in Africa were in desert margins (e.g., Sahel where rainfed agriculture in limited to a very short season when anything will grow), rainforests (e.g., slash and burn agriculture), cropland fallows (e.g., whether a fallow is 1 year or 5-year or permanent). These and numerous other issues (e.g., implementing machine learning algorithms and uncertainties inherent in them) will continue to be there in cropland mapping over such large areas as African continent. Nevertheless, advances made in this study is significant, especially in developing a nominal 30-m cropland extent of the entire continent at very good accuracies.

VII. Publications

The following publications are related to the development of the above croplands products:

1. Publications specific to this study

1) Xiong, J., Thenkabail, P. S., James C. T., Gumma, M. K., Teluguntla, P., Congalton, R. G., Poehnelt, J., Kamini Yadav., et al. (2017). A Nominal 30-m Cropland Extent of Continental Africa Using Sentinel-2 data and Landsat-8 by Integrating Random Forest (SVM) and Hierarchical Segmentation Approach on Google Earth Engine. In press.

2) Xiong, J., Thenkabail, P. S., Gumma, M. K., Teluguntla, P., Poehnelt, J., Congalton, R. G., et al. (2017). Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS Journal of Photogrammetry and Remote Sensing, 126, 225–244.

2. Peer-reviewed publications within GFSAD project

Congalton, R.G., Gu, J., Yadav, K., Thenkabail, P.S., and Ozdogan, M. 2014. Global Land Cover Mapping: A Review and Uncertainty Analysis. Remote Sensing Open Access Journal. Remote Sens. 2014, 6, 12070-12093; http://dx.doi.org/10.3390/rs61212070.

Congalton, R.G, 2015. Assessing Positional and Thematic Accuracies of Maps Generated from Remotely Sensed Data. Chapter 29, In Thenkabail, P.S., (Editor-in-Chief), 2015. "Remote Sensing Handbook" Volume I: Volume I: Data Characterization, Classification, and Accuracies: Advances of Last 50 Years and a Vision for the Future. Taylor and Francis Inc.\CRC Press, Boca Raton, London, New York. Pp. 900+. In Thenkabail, P.S., (Editor-in-Chief), 2015. "Remote Sensing Handbook" Volume I: **): Remotely Sensed Data Characterization, Classification, and Accuracies**. Taylor and Francis Inc.\CRC Press, Boca Raton, London, New York. ISBN 9781482217865 - CAT# K22125. Print ISBN: 978-1-4822-1786-5; eBook ISBN: 978-1-4822-1787-2. Pp. 678.

Gumma, M.K., Thenkabail, P.S.,Teluguntla, P., Rao, M.N., Mohammed, I.A., and Whitbread, A.M. 2016. Mapping rice-fallow cropland areas for short-season grain legumes intensification in South Asia using MODIS 250 m time-series data. International Journal of Digital Earth, http://dx.doi.org/10.1080/17538947.2016.1168489

Massey, R., Sankey, T.T., Congalton, R.G., Yadav, K., Thenkabail, P.S., Ozdogan, M., Sánchez Meador, A.J. 2017. MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types, Remote Sensing of Environment, Volume 198, 1 September 2017, Pages 490-503, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2017.06.033.

Phalke, A. R., Ozdogan, M., Thenkabail, P. S., Congalton, R. G., Yadav, K., & Massey, R. et al. (2017). A Nominal 30-m Cropland Extent and Areas of Europe, Middle-east, Russia and Central Asia for the Year 2015 by Landsat Data using Random Forest Algorithms on Google Earth Engine Cloud. (in preparation).

Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Congalton, R.G., Oliphant, A., Poehnelt, J., Yadav, K., Rao, M., and Massey, R. 2017. Spectral matching techniques (SMTs) and automated cropland classification algorithms (ACCAs) for mapping croplands of Australia using MODIS 250-m time-series (2000–2015) data, International Journal of Digital Earth.
DOI:10.1080/17538947.2016.1267269.IP-074181, http://dx.doi.org/10.1080/17538947.2016.1267269.

Teluguntla, P., Thenkabail, P., Xiong, J., Gumma, M.K., Giri, C., Milesi, C., Ozdogan, M., Congalton, R., Yadav, K., 2015. CHAPTER 6 - Global Food Security Support Analysis Data at Nominal 1 km (GFSAD1km) Derived from Remote Sensing in Support of Food Security in the Twenty-First Century: Current Achievements and Future Possibilities, in: Thenkabail, P.S. (Ed.), Remote Sensing Handbook (Volume II): Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. CRC Press, Boca Raton, London, New York., pp. 131–160. Link.

Xiong, J., Thenkabail, P.S., Tilton, J.C., Gumma, M.K., Teluguntla, P., Oliphant, A., Congalton, R.G., Yadav, K. 2017. A Nominal 30-m Cropland Extent and Areas of Continental Africa for the Year 2015 by Integrating Sentinel-2 and Landsat-8 Data using Random Forest, Support Vector Machines and Hierarchical Segmentation Algorithms on Google Earth Engine Cloud. Remote Sensing Open Access Journal (in review).

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., Thau, D. 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 126, April 2017, Pages 225-244, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2017.01.019.

3. Web sites and Data portals:

http://croplands.org (30-m global croplands visualization tool)
http://geography.wr.usgs.gov/science/croplands/index.html (GFSAD30 web portal and dissemination)
http://geography.wr.usgs.gov/science/croplands/products.html#LPDAAC (dissemination on LP DAAC)
http://geography.wr.usgs.gov/science/croplands/products.html (global croplands on Google Earth Engine)
croplands.org (crowdsourcing global croplands data)

4. Other relevant past publications prior to GFSAD project

Biggs, T., Thenkabail, P.S., Krishna, M., GangadharaRao Rao, P., and Turral, H., 2006. Vegetation phenology and irrigated area mapping using combined MODIS time-series, ground surveys, and agricultural census data in Krishna River Basin, India. International Journal of Remote Sensing. 27(19):4245-4266.

Biradar, C.M., Thenkabail, P.S., Noojipady, P., Yuanjie, L., Dheeravath, V., Velpuri, M., Turral, H., Gumma, M.K., Reddy, O.G.P., Xueliang, L. C., Schull, M.A., Alankara, R.D., Gunasinghe, S., Mohideen, S., Xiao, X. 2009. A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. International Journal of Applied Earth Observation and Geoinformation. 11(2). 114-129. doi:10.1016/j.jag.2008.11.002. January, 2009.

Dheeravath, V., Thenkabail, P.S., Chandrakantha, G, Noojipady, P., Biradar, C.B., Turral. H., Gumma, M.1, Reddy, G.P.O., Velpuri, M. 2010. Irrigated areas of India derived using MODIS 500m data for years 2001-2003. ISPRS Journal of Photogrammetry and Remote Sensing. http://dx.doi.org/10.1016/j.isprsjprs.2009.08.004. 65(1): 42-59.

Thenkabail, P.S. 2012. Special Issue Foreword. Global Croplands special issue for the August 2012 special issue for Photogrammetric Engineering and Remote Sensing. PE&RS. 78(8): 787- 788.Thenkabail, P.S. 2012. Guest Editor for Global Croplands Special Issue. Photogrammetric Engineering and Remote Sensing. PE&RS. 78(8).

Thenkabail, P.S., Biradar C.M., Noojipady, P., Cai, X.L., Dheeravath, V., Li, Y.J., Velpuri, M., Gumma, M., Pandey, S. 2007a. Sub-pixel irrigated area calculation methods. Sensors Journal (special issue: Remote Sensing

of Natural Resources and the Environment (Remote Sensing SensorsEdited by Assefa M. Melesse). 7:2519-2538. http://www.mdpi.org/sensors/papers/s7112519.pdf.

Thenkabail, P.S., Biradar C.M., Noojipady, P., Dheeravath, V., Li, Y.J., Velpuri, M., Gumma, M., Reddy, G.P.O., Turral, H., Cai, X. L., Vithanage, J., Schull, M., and Dutta, R. 2009a. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. International Journal of Remote Sensing. 30(14): 3679-3733. July, 20, 2009.

Thenkabail, P.S., Biradar, C.M., Turral, H., Noojipady, P., Li, Y.J., Vithanage, J., Dheeravath, V., Velpuri, M., Schull M., Cai, X. L., Dutta, R. 2006. An Irrigated Area Map of the World (1999) derived from Remote Sensing. Research Report # 105. International Water Management Institute. Pp. 74. Also, see under documents in: http://www.iwmigiam.org.

Thenkabail, P. S.; Dheeravath, V.; Biradar, C. M.; Gangalakunta, O. P.; Noojipady, P.; Gurappa, C.; Velpuri, M.; Gumma, M.; Li, Y. 2009b. Irrigated Area Maps and Statistics of India Using Remote Sensing and National Statistics. Journal Remote Sensing. 1:50-67. http://www.mdpi.com/2072-4292/1/2/50.

Thenkabail, P.S., GangadharaRao, P., Biggs, T., Krishna, M., and Turral, H., 2007b. Spectral Matching Techniques to Determine Historical Land use/Land cover (LULC) and Irrigated Areas using Time-series AVHRR Pathfinder Datasets in the Krishna River Basin, India. Photogrammetric Engineering and Remote Sensing. 73(9): 1029-1040. (Second Place Recipients of the 2008 John I. Davidson ASPRS President's Award for Practical papers).

Thenkabail, P.S., Hanjra, M.A., Dheeravath, V., Gumma, M.K. 2010. A Holistic View of Global Croplands and Their Water Use for Ensuring Global Food Security in the 21st Century through Advanced Remote Sensing and Non-remote Sensing Approaches. Remote Sensing open access journal. 2(1):211-261. doi:10.3390/rs2010211. http://www.mdpi.com/2072-4292/2/1/211

Thenkabail P.S., Knox J.W., Ozdogan, M., Gumma, M.K., Congalton, R.G., Wu, Z., Milesi, C., Finkral, A., Marshall, M., Mariotto, I., You, S. Giri, C. and Nagler, P. 2012. Assessing future risks to agricultural productivity, water resources and food security: how can remote sensing help? Photogrammetric Engineering and Remote Sensing, August 2012 Special Issue on Global Croplands: Highlight Article. 78(8): 773-782.

Thenkabail, P.S., Schull, M., Turral, H. 2005. Ganges and Indus River Basin Land Use/Land Cover (LULC) and Irrigated Area Mapping using Continuous Streams of MODIS Data. Remote Sensing of Environment. Remote Sensing of Environment, 95(3): 317-341.

Velpuri, M., Thenkabail, P.S., Gumma, M.K., Biradar, C.B., Dheeravath, V., Noojipady, P., Yuanjie, L.,2009. Influence of Resolution or Scale in Irrigated Area Mapping and Area Estimations. Photogrammetric Engineering and Remote Sensing (PE&RS). 75(12): December 2009 issue.

5. Books and Book Chapters

Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Giri, C., Milesi, C., Ozdogan, M., Congalton, R.,Tilton, J.,Sankey,T.R., Massey, R., Phalke, A., and Yadav, K. 2015. Global Food Security Support Analysis Data at Nominal 1 km (GFSAD1 km) Derived from Remote Sensing in Support of Food Security in the Twenty-First Century: Current Achievements and Future Possibilities, Chapter 6. In Thenkabail, P.S., (Editor-in-Chief), 2015. "Remote Sensing Handbook" (Volume II): Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. Taylor and Francis Inc. Press, Boca Raton, London, New York. ISBN 9781482217957 - CAT# K22130. Pp. 131-160

Biradar, C.M., Thenkabail. P.S., Noojipady, P., Li, Y.J., Dheeravath, V., Velpuri, M., Turral, H., Cai, X.L., Gumma, M., Gangalakunta, O.R.P., Schull, M., Alankara, R.D., Gunasinghe, S., and Xiao, X. 2009. Book Chapter 15: Global map of rainfed cropland areas (GMRCA) and stastistics using remote sensing. Pp. 357-392. In the book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 475. Published in June, 2009. (Editors: Thenkabail. P., Lyon, G.J., Biradar, C.M., and Turral, H.).

Gangalakunta, O.R.P., Dheeravath, V., Thenkabail, P.S., Chandrakantha, G., Biradar, C.M., Noojipady, P., Velpuri, M., and Kumar, M.A. 2009. Book Chapter 5: Irrigated areas of India derived from satellite sensors and national statistics: A way forward from GIAM experience. Pp. 139-176. In the book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 475. Published in June, 2009. (Editors: Thenkabail. P., Lyon, G.J., Biradar, C.M., and Turral, H.).

Li, Y.J., Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Velpuri, M., Gangalakunta, O.R., Cai, X.L. 2009. Book Chapter 2: A history of irrigated areas of the world. Pp. 13-40. In the book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 475. Published in June, 2009. (Editors: Thenkabail. P.,Lyon, G.J., Biradar, C.M., and Turral, H.).

Thenkabail, P.S., Lyon, G.J., and Huete, A. 2011. Book Chapter # 1: Advances in Hyperspectral Remote Sensing of Vegetation. In Book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Edited by Thenkabail, P.S., Lyon, G.J., and Huete, A. Pp. 3-38.

Thenkabail. P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Gumma, M., Li, Y.J., Velpuri, M., Gangalakunta, O.R.P. 2009c. Book Chapter 3: Global irrigated area maps (GIAM) and statistics using remote sensing. Pp. 41-120. In the book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 475. Published in June, 2009. (Editors: Thenkabail. P., Lyon, G.J., Biradar, C.M., and Turral, H.).

Thenkabail. P., Lyon, G.J., Turral, H., and Biradar, C.M. (Editors) 2009d. Book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 556 (48 pages in color). Published in June, 2009. Reviews of this book: http://www.crcpress.com/product/isbn/9781420090093 http://gfmt.blogspot.com/2011/05/review-remote-sensing-of-global.html

Thenkabail, P.S. and Lyon, J.G. 2009. Book Chapter 20: Remote sensing of global croplands for food security: way forward. Pp. 461-466. In the book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 475. Published in June, 2009. (Editors: Thenkabail. P., Lyon, G.J., Biradar, C.M., and Turral, H.).

Turral, H., Thenkabail, P.S., Lyon, J.G., and Biradar, C.M. 2009. Book Chapter 1: Context, need: The need and scope for mapping global irrigated and rain-fed areas. Pp. 3-12. In the book entitled: "Remote Sensing of Global Croplands for Food Security" (CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 475. Published in June, 2009. (Editors: Thenkabail. P., Lyon, G.J., Biradar, C.M., and Turral, H.).

VIII. Acknowledgements

The project was funded by the National Aeronautics and Space Administration (NASA) grant number: NNH13AV82I through its MEaSUREs (Making Earth System Data Records for Use in Research Environments) initiative. The United States Geological Survey (USGS) provided supplemental funding from other direct and indirect means through the Climate and Land Use Change Mission Area, including the Land Change Science (LCS) and Land Remote Sensing (LRS) programs. The project was led by United States Geological Survey (USGS) in collaboration with NASA AMES, University of New Hampshire (UNH), California State University Monterey Bay (CSUMB), University of Wisconsin (UW), NASA GSFC, and Northern Arizona University. There were a number of International partners including The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Authors gratefully acknowledge the excellent support and guidance received from the LP DAAC team members (Carolyn Gacke, Lindsey Harriman, Sydney Neeley), as well as Chris Doescher, LP DAAC project manager when releasing these data. We also like to thank Susan Benjamin, Director of USGS Western Geographic Science Center (WGSC) as well as WGSC administrative officer Larry Gaffney for their cheerful support and encouragement throughout the project.

IX. Contact Information

LP DAAC User Services
U.S. Geological Survey (USGS)
Center for Earth Resources Observation and Science (EROS)
47914 252nd Street
Sioux Falls, SD 57198-0001

Phone Number: 605-594-6116
Toll Free: 866-573-3222 (866-LPE-DAAC)
Fax: 605-594-6963

Email: lpdaac@usgs.gov
Web: https://lpdaac.usgs.gov

For the Principal Investigators, feel free to write to:
Prasad S. Thenkabail at pthenkabail@usgs.gov

For 30-m cropland extent product of Africa, please contact:
Jun Xiong at jxiong@usgs.gov, jun.xiong1981@gmail.com
Prasad S. Thenkabail at pthenkabail@usgs.gov
Pardhasaradhi Teluguntla at pteluguntla@usgs.gov
More details about the GFSAD project and products can be found at: globalcroplands.org

X. Citations

Xiong, J., Thenkabail, P.S., Tilton, J.C., Gumma, M.K., Teluguntla, P., Congalton, R.G., Yadav, K., Dungan, J., Oliphant, A.J., Poehnelt, J., Smith, C., Massey, R. (2017). *NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) Cropland Extent 2015 Africa 30 m V001* [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MEaSUREs/GFSAD/GFSAD30AFCE.001

XI. References

Barazzetti, L., Cuca, B., Previtali, M., 2016. Evaluation of registration accuracy between Sentinel-2 and Landsat 8, in: Themistocleous, K., Hadjimitsis, D.G., Michaelides, S., Papadavid, G. (Eds.), Fourth International Conference on Remote Sensing and Geoinformation of the Environment. SPIE, pp. 968809–968809–9.

Büttner, G., 2014. CORINE Land Cover and Land Cover Change Products, in: Land Use and Land Cover Mapping in Europe. Springer Netherlands, Dordrecht, pp. 55–74.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Zhang, W., Lu, M., Tong, X., Mills, J., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing 103, 7.

Congalton, R.. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment. Vol. 37, pp. 35-46.

Congalton, R. and K. Green. 2009. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. 2nd Edition. CRC/Taylor & Francis, Boca Raton, FL 183p

Costa, H., Carrão, H., Bação, F., Caetano, M., 2014. Combining per-pixel and object-based classifications for mapping land cover over large areas. International Journal of Remote Sensing 35, 738–753.

Dingle Robertson, L., King, D.J., 2011. Comparison of pixel- and object-based classification in land cover change mapping. International Journal of Remote Sensing 32, 1505–1529.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment 120, 25–36.

D'Odorico, P., Gonsamo, A., Damm, A., Schaepman, M.E., 2013. Experimental Evaluation of Sentinel-2 Spectral Response Functions for NDVI Time-Series Continuity. IEEE Transactions on Geoscience and Remote Sensing 51, 1336–1348.

Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sensing of Environment 103, 179–189.

Gueymard, C. (2001). Parameterized transmittance model for direct beam and circumsolar spectral irradiance. Solar Energy, 71(5):325–346.

Haack, B., Mahabir, R., Kerkering, J., 2014. Remote sensing-derived national land cover land use maps: a comparison for Malawi. Geocarto International 30, 270–292.

Hentze, K., Thonfeld, F., Menz, G., 2016. Evaluating Crop Area Mapping from MODIS Time-Series as an Assessment Tool for Zimbabwes Fast Track Land Reform Programme. PLoS ONE 11, e0156630.

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 Sensing 8, 666.

Huang, C., Davis, L.S., Townshend, J.R.G., 2010. An assessment of Support Vector Machines for land cover classification. International Journal of Remote Sensing 23, 725–749.

Im, J., Jensen, J.R., Tullis, J.A., 2008. Object-based change detection using correlation image analysis and image segmentation. International Journal of Remote Sensing 29, 399–423.

Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: The Landsat Data Continuity Mission. Remote Sensing of Environment 122, 11–21.

Kidane, Y., Stahlmann, R., Beierkuhnlein, C., 2012. Vegetation dynamics, and land use and land cover change in the Bale Mountains, Ethiopia. Environmental Monitoring and Assessment 184, 7473–7489.

Kruger, A.C., 2006. Observed trends in daily precipitation indices in South Africa: 19102004. International Journal of Climatology 26, 2275–2285.

Lambert, M.-J., Waldner, F., Defourny, P., 2016. Cropland Mapping over Sahelian and Sudanian Agrosystems: A Knowledge-Based Approach Using PROBA-V Time Series at 100-m. Remote Sensing 8, 232.

Languille, F., Déchoz, C., Gaudel, A., Greslou, D., Lussy, F. de, Trémas, T., Poulain, V., 2015. Sentinel-2 geometric image quality commissioning: first results, in: Bruzzone, L. (Ed.), SPIE Remote Sensing. SPIE, pp. 964306–964306–13.

Latifovic, R., Homer, C., Ressl, R., Pouliot, D., Hossain, S.N., Colditz Colditz, R.R., Olthof, I., Giri, C. and Victoria, A., 2010. North American land Change Monitoring System (NALCMS). Remote sensing of land use and land cover: principles and applications. CRC Press, Boca Raton.

Lillesand, T., Kiefer, R.W. and Chipman, J., 2014. Remote sensing and image interpretation. John Wiley & Sons.
Löw, F., Schorcht, G., Michel, U., Dech, S., Conrad, C., 2012. Per-field crop classification in irrigated agricultural regions in middle Asia using Random Forest and support vector machine ensemble, in: Michel, U., Civco, D.L., Ehlers, M., Schulz, K., Nikolakopoulos, K.G., Habib, S., Messinger, D., Maltese, A. (Eds.), SPIE Remote Sensing. SPIE, pp. 85380R–85380R–11.

Malinverni, E.S., Tassetti, A.N., Mancini, A., Zingaretti, P., Frontoni, E., Bernardini, A., 2011. Hybrid object-based approach for land use/land cover mapping using high spatial resolution imagery. International Journal of Geographical Information Science 25, 1025–1043.

Motha, R.P., Leduc, S.K., Steyaert, L.T., Sakamoto, C.M., Strommen, N.D., Motha, R.P., Leduc, S.K., Steyaert, L.T., Sakamoto, C.M., Strommen, N.D., 1980. Precipitation Patterns in West Africa. http://dx.doi.org/10.1175/1520-0493(1980)108<1567:PPIWA>2.0.CO;2 108, 1567–1578.

Myint, S.W., Myint, S.W., Gober, P., Brazel, A., Gober, P., Brazel, A., Grossman-Clarke, S., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment 115, 1145–1161.

Pal, M., 2007. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. International Journal of Remote Sensing 27, 2877–2894.

Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. Remote Sensing of Environment 187, 156–168.

Rembold, F., Carnicelli, S., Nori, M., Ferrari, G.A., 2000. Use of aerial photographs, Landsat TM imagery and multidisciplinary field survey for land-cover change analysis in the lakes region (Ethiopia). International Journal of Applied Earth Observation and Geoinformation 2, 181–189.

Roy, D.P., Wulder, M.A., Loveland, T.R., C E, W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S.,

Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z., 2014. Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment 145, 154–172.

Shalaby, A., Tateishi, R., 2007. Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. Applied Geography 27, 28–41.

Shi, D., Yang, X., 2015. Support Vector Machines for Land Cover Mapping from Remote Sensor Imagery, in: Monitoring and Modeling of Global Changes: A Geomatics Perspective. Springer Netherlands, Dordrecht, pp. 265–279.

Storey, J., Roy, D.P., Masek, J., Gascon, F., Dwyer, J., Choate, M., 2016. A note on the temporary misregistration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery. Remote Sensing of Environment 186, 121–122.

Story, M. and R. Congalton. 1986. Accuracy assessment: A user's perspective. Photogrammetric Engineering and Remote Sensing. Vol. 52, No. 3. pp. 397-399.
Stow, D., Hamada, Y., Coulter, L., Anguelova, Z., 2008. Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. Remote Sensing of Environment 112, 1051–1061.

Sulla-Menashe, D., Friedl, M.A., Krankina, O.N., Baccini, A., Woodcock, C.E., Sibley, A., Sun, G., Kharuk, V., Elsakov, V., 2011. Hierarchical mapping of Northern Eurasian land cover using MODIS data. Remote Sensing of Environment 115, 392–403.

Tateishi, R., Hoan, N.T., Kobayashi, T., 2014. Production of Global Land Cover DataGLCNMO2008. Journal of Geography and Geology 6, 99.

Teluguntla, P., Thenkabail, P., Xiong, J., Gumma, M.K., Giri, C., Milesi, C., Ozdogan, M., Congalton, R., Yadav, K., 2015. CHAPTER 6 - Global Food Security Support Analysis Data at Nominal 1 km (GFSAD1km) Derived from Remote Sensing in Support of Food Security in the Twenty-First Century: Current Achievements and Future Possibilities, in: Thenkabail, P.S. (Ed.), Remote Sensing Handbook(Volume Ii):Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. CRC Press, Boca Raton, London, New York., pp. 131–160.

Thenkabail, P.S., Hanjra, M. a, Dheeravath, V., Gumma, M., 2010. A Holistic View of Global Croplands and Their Water Use for Ensuring Global Food Security in the 21st Century through Advanced Remote Sensing and Non-remote Sensing Approaches. Remote Sensing 2, 211.

Tian, S., Zhang, X., Tian, J., Sun, Q., 2016. Random Forest Classification of Wetland Landcovers from Multi-Sensor Data in the Arid Region of Xinjiang, China. Remote Sensing 8, 954.

Tilton, J.C., Tarabalka, Y., Montesano, P.M., Gofman, E., 2012. Best Merge Region-Growing Segmentation With Integrated Nonadjacent Region Object Aggregation. IEEE Transactions on Geoscience and Remote Sensing 50, 4454–4467.

Tilton, J. C, Pasolli, E. 2014. Incorporating Edge Information into Best Merge Region Growing Segmentation. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Quebec, Canada, pp. 4891-4894.

Vapnik, V.N., Vapnik, V., 1998. Statistical learning theory.

Waldner, F., De Abelleyra, D., Verón, S.R., Zhang, M., Wu, B., Plotnikov, D., Bartalev, S., Lavreniuk, M., Skakun, S., Kussul, N., Le Maire, G., Dupuy, S., Jarvis, I., Defourny, P., 2016. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity. International Journal of Remote Sensing 37, 3196–3231.

Wang, Q., Blackburn, G.A., Onojeghuo, A.O., Dash, J., Zhou, L., Zhang, Y., and Atkinson. P.M. 2017. Fusion of Landsat 8 OLI and Sentinel-2 MSI Data. Available from: https://www.researchgate.net/publication/316028249_Fusion_of_Landsat_8_OLI_and_Sentinel-2_MSI_Data [accessed Sep 7, 2017]. Digital Object Identifier 10.1109/TGRS.2017.2683444

Were, K.O., Dick, Ø.B., Singh, B.R., 2013. Remotely sensing the spatial and temporal land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya. Applied Geography 41, 75–86.

Werff, H. van der Meer, , F., 2016. Sentinel-2A MSI and Landsat 8 OLI Provide Data Continuity for Geological Remote Sensing. Remote Sensing 8, 883.

Yu, L., Wang, J., Gong, P., 2013. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach. International Journal of Remote Sensing.

Zhao, Y., Gong, P., Yu, L., Hu, L., Li, X., Li, C., Zhang, H., Zheng, Y., Wang, J., Zhao, Y., Cheng, Q., Liu, C., Liu, S., Wang, X., 2014. Towards a common validation sample set for global land-cover mapping. International Journal of Remote Sensing 35, 4795–4814.

Zucca, C., Wu, W., Dessena, L., Mulas, M., 2015. Assessing the Effectiveness of Land Restoration Interventions in Dry Lands by Multitemporal Remote Sensing A Case Study in Ouled DLIM (Marrakech, Morocco). Land Degradation & Development 26, 80–91.