Article

Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine

Jun Xiong [1,2,*] ![ID], **Prasad S. Thenkabail** [1], **James C. Tilton** [3], **Murali K. Gumma** [4] ![ID], **Pardhasaradhi Teluguntla** [1,2] ![ID], **Adam Oliphant** [1] ![ID], **Russell G. Congalton** [5] ![ID], **Kamini Yadav** [5] ![ID] **and Noel Gorelick** [6]

[1] Western Geographic Science Center, U. S. Geological Survey (USGS), 2255, N. Gemini Drive, Flagstaff, AZ 86001, USA; pthenkabail@usgs.gov (P.S.T.); pteluguntla@usgs.gov (P.T.); aoliphant@usgs.gov (A.O.)
[2] Bay Area Environmental Research Institute (BAERI), 596 1st St West Sonoma, CA 95476, USA
[3] Computational & Information Science and Technology Office, Mail Code 606.3, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; james.c.tilton@nasa.gov
[4] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad 502324, India; m.gumma@cgiar.org
[5] Department of Natural Resources and the Environment, University of New Hampshire, 56 College Road, Durham, NH 03824, USA; russ.congalton@unh.edu (R.G.C.); kaminiyadav.02@gmail.com (K.Y.)
[6] Google Inc., 1600 Amphitheater Parkway, Mountain View, CA 94043, USA; gorelick@google.com
*   Correspondence: jun.xiong1981@gmail.com; Tel.: +1-928-556-7215

Abstract: A satellite-derived cropland extent map at high spatial resolution (30-m or better) is a must for food and water security analysis. Precise and accurate global cropland extent maps, indicating cropland and non-cropland areas, are starting points to develop higher-level products such as crop watering methods (irrigated or rainfed), cropping intensities (e.g., single, double, or continuous cropping), crop types, cropland fallows, as well as for assessment of cropland productivity (productivity per unit of land), and crop water productivity (productivity per unit of water). Uncertainties associated with the cropland extent map have cascading effects on all higher-level cropland products. However, precise and accurate cropland extent maps at high spatial resolution over large areas (e.g., continents or the globe) are challenging to produce due to the small-holder dominant agricultural systems like those found in most of Africa and Asia. Cloud-based geospatial computing platforms and multi-date, multi-sensor satellite image inventories on Google Earth Engine offer opportunities for mapping croplands with precision and accuracy over large areas that satisfy the requirements of broad range of applications. Such maps are expected to provide highly significant improvements compared to existing products, which tend to be coarser in resolution, and often fail to capture fragmented small-holder farms especially in regions with high dynamic change within and across years. To overcome these limitations, in this research we present an approach for cropland extent mapping at high spatial resolution (30-m or better) using the 10-day, 10 to 20-m, Sentinel-2 data in combination with 16-day, 30-m, Landsat-8 data on Google Earth Engine (GEE). First, nominal 30-m resolution satellite imagery composites were created from 36,924 scenes of Sentinel-2 and Landsat-8 images for the entire African continent in 2015–2016. These composites were generated using a median-mosaic of five bands (blue, green, red, near-infrared, NDVI) during each of the two periods (period 1: January–June 2016 and period 2: July–December 2015) plus a 30-m slope layer derived from the Shuttle Radar Topographic Mission (SRTM) elevation dataset. Second, we selected Cropland/Non-cropland training samples (sample size = 9791) from various sources in GEE to create pixel-based classifications. As supervised classification algorithm, Random Forest (RF) was used as the primary classifier because of its efficiency, and when over-fitting issues of

RF happened due to the noise of input training data, Support Vector Machine (SVM) was applied to compensate for such defects in specific areas. Third, the Recursive Hierarchical Segmentation (RHSeg) algorithm was employed to generate an object-oriented segmentation layer based on spectral and spatial properties from the same input data. This layer was merged with the pixel-based classification to improve segmentation accuracy. Accuracies of the merged 30-m crop extent product were computed using an error matrix approach in which 1754 independent validation samples were used. In addition, a comparison was performed with other available cropland maps as well as with LULC maps to show spatial similarity. Finally, the cropland area results derived from the map were compared with UN FAO statistics. The independent accuracy assessment showed a weighted overall accuracy of 94%, with a producer's accuracy of 85.9% (or omission error of 14.1%), and user's accuracy of 68.5% (commission error of 31.5%) for the cropland class. The total net cropland area (TNCA) of Africa was estimated as 313 Mha for the nominal year 2015. The online product, referred to as the Global Food Security-support Analysis Data @ 30-m for the African Continent, Cropland Extent product (GFSAD30AFCE) is distributed through the NASA's Land Processes Distributed Active Archive Center (LP DAAC) as (available for download by 10 November 2017 or earlier): https://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30AFCE.001 and can be viewed at https://croplands.org/app/map. Causes of uncertainty and limitations within the crop extent product are discussed in detail.

---

1. Introduction

Agricultural areas are changing rapidly over time and space across the world as a result of land cover change as well as climate variability. Mapping the geographical extent of croplands, their precise locations, and establishing areas of agricultural croplands is of great importance for managing food production systems and to study their inter-relationships with geo-political, socio-economic, health, environmental, and ecological issues [1]. In many food-insecure regions of the world, such as Africa, understanding and characterizing agricultural production remains a major challenge [2]. In addition, a primary requirement for agricultural cropland studies is a dependency on the availability of a precise map of cropland extent at high spatial resolution (30-m or better) as well as determining reliable and consistent cropland areas derived from these accurate maps [3]. The absence of such a product leads to great uncertainties in all higher-level cropland products resulting in poor assessment of global and local food security scenarios. Consequently, the demand for a baseline cropland extent product at high resolution and accuracy has been widely recognized [4]. Accuracy of higher-level cropland products such as cropping intensities, crop types, crop watering methods (e.g., irrigated or rainfed), planted or left fallow, crop health, crop productivity (productivity per unit of land, $kg \cdot m^{-2}$), and crop water productivity (productivity per unit of water or crop per drop, $kg \cdot m^{-3}$) are dependent on having a precise cropland extent product as a baseline product. In Africa, these products are particularly helpful due to the absence of high resolution cropland products that map field level details of croplands making them an invaluable baseline product for all higher-level products such as crop type, crop productivity, and crop water productivity [5,6].

Remote sensing has long been recognized as an effective tool for broad-scale crop mapping [7–9]. The two most applied remote sensing methods for land-cover mapping are manual classification based on visual interpretation [10] and digital per-pixel classification [11]. Although the human capacity for interpreting images is remarkable, visual interpretation is subjective, time-consuming, and expensive on large area. A number of cropland cover datasets on a global scale have been developed, mostly at a coarse resolution of 1-km [8,12–14]. Others have mapped cropland as

one class in their land cover products at MODIS resolution [15–19]. However, all these studies suffered from inability to depict individual farm fields. Moving from existing products to high resolution (30-m or finer) greatly improves the ability to capture small and fragmented cultivated fields. On this topic, National Agricultural Statistics Service (NASS) of the US Department of Agriculture (USDA) produced Cropland Data Layers (CDLs) for the US using a decision tree approach based on millions ground samples generated from farmers' surveys over across the country as well as the National Land Cover Database [20]. However, such advanced operational approaches cannot be replicated in developing regions other than North America and Europe because of the lack of systematic collection of ground training samples. Alternate procedures have consisted of unsupervised approaches [13,21–23] and supervised methods in small regional areas with different classifiers including decision trees [12], Support Vector Machine [24,25], Random Forest [12], neural networks [26–28], data mining [29], and hybrid methods [30]. In order to improve classification results, the following issues were investigated in literature which include the selection of the dates [31], temporal windows derivation [32], input features selection [33] and automated classification methods [34]. Object-based approaches of crop identification have also been explored [35].

One issue that has confounded above cropland mapping efforts is how the term "cropland" has been defined. For instance, the U.S. Department of Agriculture (USDA) includes in its cropland definition "all areas used for the production of adapted crops for harvest", which covers both cultivated and non-cultivated areas [36]. In most global land cover products, such as Africover [37], GLC2000 [38], GlobCover [39], GLCShar [40], MODIS Land Cover [41] croplands are partly combined in mosaic or mixed classes including meadows and pastures [42], making them difficult to use in agricultural applications, either as agricultural masks or as a source for area estimates. In a previous effort to compile four existing global cropland into a 1-km global cropland extent map [43], cropland was defined as: "lands cultivated with plants harvested for food, feed, and fiber, include both seasonal crops (e.g., wheat, rice, corn, soybeans, cotton) and continuous plantations (e.g., coffee, tea, rubber, cocoa, oil palms). Cropland fallows are lands uncultivated during a season or a year but are farmlands that are equipped for cultivation, and hence included as part of croplands. From a remote-sensing perspective, a cropland in this study is a piece of land of minimum 0.09 ha (30-m × 30-m pixel) that is sowed/planted and harvest-able at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and can sometimes be combined with some tree or woody vegetation. Some crops like sugarcane plantation and cassava crop are not necessarily planted yearly, but are still crops based on planting that had taken place during a previous year. Greenhouses, and aquaculture are part of the farmlands and have different signature from other croplands [44], but these are negligible in Africa. In a nutshell, the cropland extent includes: standing crops, cropland fallows and plantations.

In order to make continental scale classification feasible, new methods and approaches need to be adopted or developed to deal with the complex classification issues [45–48]. In this research, we propose this integrated method of pixel-based classification and object-based segmentation for large area cropland mapping. A number of earlier studies [49–52] have explored such integrated approaches. Pixel-based classification algorithms, such as the Random Forests (RF) and Support Vector Machines (SVM), are widely used due to their efficiency over large areas. These pixel-based clustering algorithms focus only on the spectral value of each pixel and often result in image speckle and overall inaccuracies when applied to high resolution imagery. Since each pixel is dealt with in isolation from its neighbors in the pixel-based paradigm, close neighbors often have different classes, despite being similar. When classification to produce discrete mapped entities is needed, an object-based segmentation approach can alleviate such problems [53]. For object-based classification, field boundaries can be derived either from a digital vector database [54] or by segmentation [55]. In landscapes with mixed agriculture and pastoral land cover classes (e.g., Sahelian countries), image segmentation methods seem to provide a considerable advantage, since these land cover types are structurally fairly dissimilar to non-cropland areas whereas they are spectrally similar [56].

The aim of this paper is to develop a 30-m crop extent map of continental Africa by integrating pixel-based and object-based approaches. The generic methodology is capable of handling high-resolution satellite imagery with support of cloud-based Google Earth Engine (GEE) computing platform [57]. First, creating two half-year period mosaics from 30-m Landsat-8 Operational Land Imager (OLI) Top of Atmosphere (TOA) and 10-m to 20-m Sentinel-2 multi-spectral instrument (MSI) L1C product from 2015 to 2016; Second, two pixel-based supervised classifiers (Random Forest and Support Vector Machines) are applied to input dataset to obtain pixel-based classification; Third, object-based segmentations from Recursive Hierarchical Image Segmentation (RHSeg) were introduced to improve the pixel-based classification. Fourth, the study will compare the cropland areas determined using 30-m cropland product for the nominal 2015 with other statistics such as from the Food and Agricultural Organization (FAO) of the United Nations (UN). This research is a part of the Global Food Security-Support Analysis Data Project at 30-m (GFSAD30).

2. Materials

2.1. Study Area

We have chosen the entire continent of Africa for the study area (Figure 1), which extends from approximately 38° N to 35° S , occupies 3037 million hectares (Mha), and has 7 distinct geologic and bio-geographic regions with varying land cover types [58]. Demographic changes of continental Africa are expected to be staggering in the 21st Century with population expected to increase from the current 1.2 billion to over 4 billion by the end of the Century [59]. Africa is endowed with a wide diversity of agro-ecological zones. These zones range from the heavy rain-forest vegetation with bi-annual rainfall to relatively sparse, dry and arid vegetation with low unimodal rainfall. This diversity is a tremendous asset, but it also poses a substantial challenge for agricultural development. On the one hand, it creates a vast potential with respect to the mix of agricultural commodities and products which can be produced and marketed in domestic and external markets. On the other hand, the diversity implies that there are no continent-wide uniform solutions to agricultural developmental problems across the continent. Thereby, a precise and accurate cropland extent map of Africa is of significant importance to study crop dynamics, water security, and food security.

2.2. Cloud-Free Satellite Imagery Composition at 30-m Resolution

Concurrent availability of 10-day Sentinel-2 and 16-day Landsat-8 data provides an unprecedented opportunity to gather high resolution data for global cropland mapping. Sentinel-2 data become available for Africa and Europe in the middle of 2015 [60] and its capabilities to map crop types and tree species have been assessed [34,61–64] and its 10-m and 20-m data provides much more details than Landsat 30-m data. The easy and simultaneous access to entire archive of Sentinel-2 and Landsat-8 products through GEE, as well as the fast and scalable computational tools that it offers, makes GEE an essential and powerful tool for this project. Creating 30-m cloud-free imagery composition for entire Africa is a challenging task because of west and east African monsoon [65], which cause constant clouds in Gulf of Guinea and east Mozambique most of the time. As a result, we established the Sentinel-2 composite during two periods (period 1: January–June 2016, and period 2: July–December 2015), each coinciding with the two main crop growing seasons [24,44,66–69] in Africa. All data were resampled to 30-m using the average value of all involved pixels. However, data-gaps still existed in small portions of the continent due to cloud and haze issues after this composition. For these gaps, Landsat multi-bands (Table 1) with similar wavelength range as Sentinel-2 MSI were used as supplementary data for gap-filling, to make sure this 30-m wall-to-wall continental mosaic was cloud-free. In the end, a total of 36,924 images (20,214 from Sentinel-2 and 16,710 from Landsat 8) were queried from GEE data-pool and was used in this study.
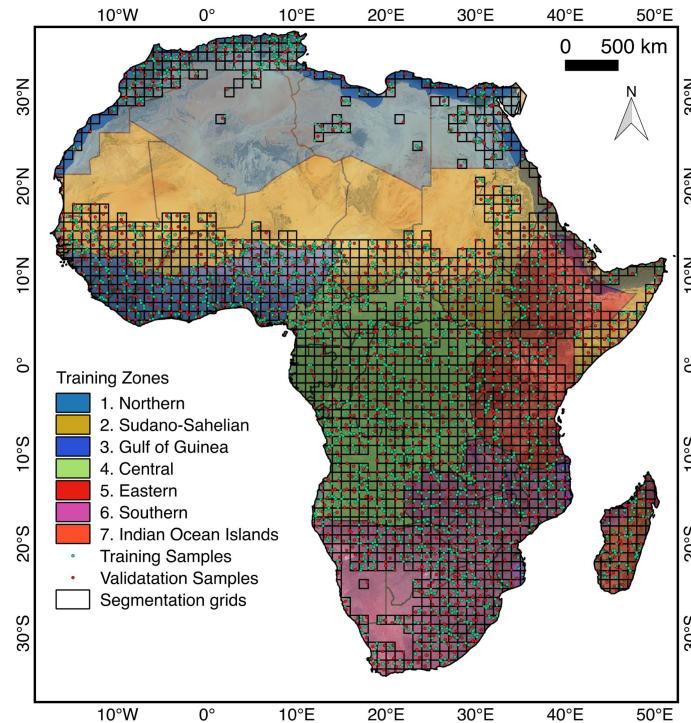
Figure 1. Map of Africa and its seven stratified zones used in the study. The pixel-based supervised classifications were run separately for each of the 7 refined agro-ecological zones (RAEZs, or simply referred to as zones: Northern, Sudano-Sahelian, Gulf of Guinea, Central, Eastern, Southern, Indian Ocean Islands). The object-based segmentation was run on each $1° \times 1°$ grids to delineate crop field boundaries. The dots represent the location of the reference training and validation samples (green indicates training and red indicates validation).

Table 1. Characteristics of Sentinel-2 MSI and Landsat-8 OLI bands used in this study. Five bands were used for each of the two periods plus slope layer from SRTM (Total 11 bands for input of classification).

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

The gap-filling of Sentinel-2 data with Landsat-8 data poses some technical challenges and requires imagery to be harmonized. The platforms and sensors differ in their orbital, spatial, and spectral configuration. As a consequence, measured physical values and radiometric attributes of the imagery are affected. For example, a root mean square error (RMSE) greater than 8% in the red band was found when comparing Sentinel and Landsat simulated data, due to the discrepancies in the nominal relative spectral response functions (RSRF) [70]. Werff compared Sentinel-2A MSI and Landsat-8 OLI Data [71], finding the correlation of their TOA reflectance products is higher than their bottom-of-atmosphere

reflectance products. Besides, the combined use of 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 an expected sensor-to-sensor misregistration of 38-m [72]. This is because although both sensor geolocation systems use parametric approaches, whereby information concerning the sensing geometry is modeled and the sensor exterior orientation parameters (altitude and position) are measured, they use different ground control and digital elevation models to refine the geolocation [72,73]. These misalignments vary geographically but should be stable for a given area. A study demonstrates that sub-pixel accuracy was achieved between 10-m resolution Sentinel-2 bands (band 3) and 15-m resolution panchromatic Landsat images (band 8) [74]. We determined that the mismatch between the geo-referencing of Landsat and Sentinel is within 30-m by comparing multiple ground control points from obvious, well-recognized locations on the land when both sensors images were available. Sentinel-2 has two NIR bands B8 (10-m) and B8A (20-m); B8 is consistently lower than B8A due to different gain settings. The B8A band was used here for 'NIR band' because it matched Landsat data better.

For each period (January–June 2016, July–December 2015) 5 bands (blue, green, red, NIR and an NDVI band (Table 1, Figure 2) were composited using Sentinel-2 and Landsat-8 combined. First, TOA reflectance values of Sentinel-2 were calculated and mosaicked using median values to create layer stacks for each of the seasons separately. Wherever "data-gaps" found, they were identified and filled using Landsat-8 data. Note that each of the 5 bands were composited over each of the two periods so we had total 10 bands from two time periods. In addition, we derived a slope surface from the Shuttle Radar Topography Mission (SRTMGL1v3) [75] digital elevation at one arc-sec (approximately 30-m) resolution dataset. These 11 bands were organized as a GEE Image Collection object, which provided a programmable way to run classification algorithms such as RF and SVM deployed on GEE.
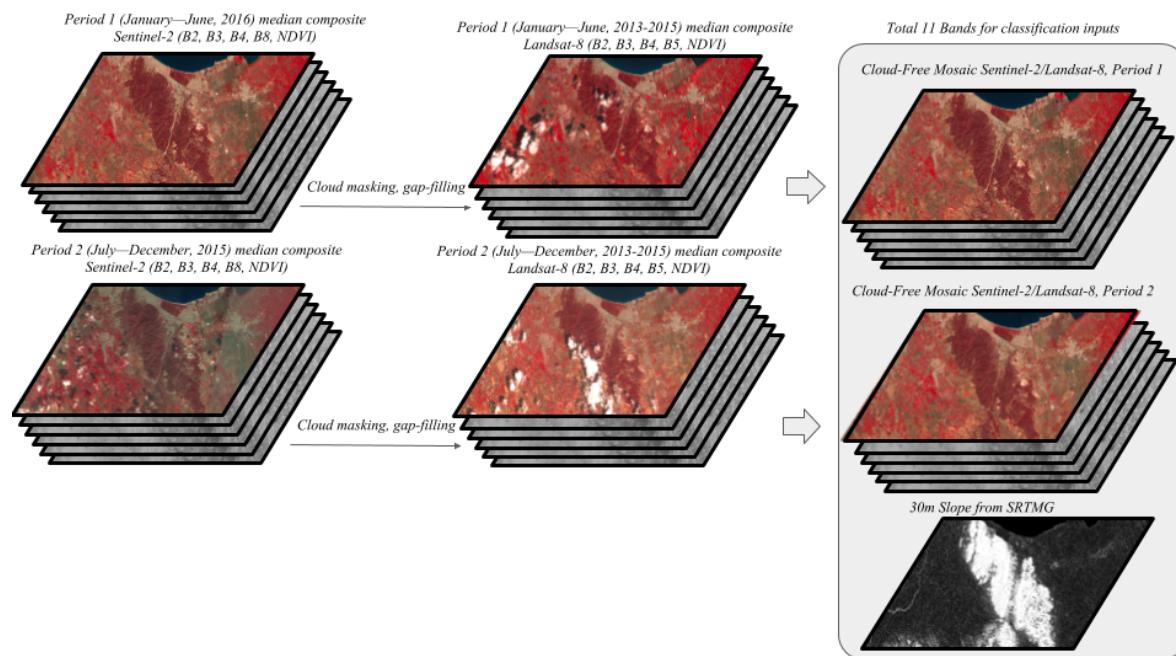

Figure 2. Illustration of the 11 input layers in this study. Five bands (Blue, Green, Red, NIR, NDVI) were composited for each of two periods (period 1: January–June 2016, and period 2: July–December 2015) for entire African continent. Total 10 bands plus the slope layer derived from SRTM elevation data were composed on GEE for pixel-based classification, 10 bands without slope layer were used for object-based segmentation.

It is noteworthy that half-yearly (January–June, July–December) composites were a measure of expediency in attaining wall-to-wall cloud-free mosaics over such a large area as Africa. Because of cloud cover, bimonthly mosaics or even trimonthly mosaics contain areas with no-data. In order to create wall-to-wall cloud-free mosaics composites of Africa, we determined through experiments that only 2 composites per year (half-yearly as mentioned above) could be reliably generated. Data availability is much higher in some areas (e.g., North Africa) than others (central Africa), which means we could set shorter composite periods in these regions, to get more composite periods within one year. However, it also makes the number of inputs bands different between regions, which increases difficulty in operational processing. Eventually, nominal 30-m cloud-free imagery were generated for the entire African continent for two main crop growing periods (January–June 2016; July–December 2015).

2.3. Reference Training Samples

We obtained reference training data (Figure 1) from following reliable sources in addition to our own collections. First, we randomly distributed 10,000 data points across land of continental Africa. We assessed these samples to ensure that they represent homogeneous cropland or non-cropland classes in a 90-m × 90-m sample frame using National Geospatial Agency (NGA) sub-meter to 5-m imagery. We removed some heterogeneous (e.g., cropland mixed with non-cropland) samples. After using 5511 randomly generated samples to train the pixel-based classification algorithm (RF and SVM), another 4280 polygons were appended which were placed by the analyst. In the end, there were 9791 training samples that were either croplands or non-croplands derived from VHRI spread across Africa. The attributes, ground photos or satellite images of these training samples are accessible through https://web.croplands.org/app/data/search.

The reference training data were then used to generate a cropland versus non-cropland knowledge-base for the algorithms. For an example of zone 1 sample sets, reflectance values of non-croplands were much higher than cropland samples (e.g., Figure 3), especially for Band 4 (red), Band 3 (green) and Band 2 (blue), while NDVI values of croplands were much higher than non-croplands. In this sample area (Figure 3), period 1 (January–June 2016) reflectance was significantly higher than period 2 (July–December 2015) due to greater intensity of cropping during this period in this sample area as well as types of crops grown during period 1 compared to period 2 (Figure 2). However, this may change in other areas depending on crop dynamics. This knowledge was used in training, classifying, and separating croplands from non-croplands in the the pixel-based supervised classification algorithms (RF, SVM).

2.4. Reference Validation Sample Polygons

Reference validation samples were also collected using similar approach as in Section 2.3. Spatial distribution of the validation data is shown in Figure 1. The validation data was hidden from the map producers and was made available only to independent accuracy assessment team. These reference datasets are publicly available for download at: https://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30AFCE.001. Accuracy error matrices were established for each of the 7 refined agro-ecological zones or RAEZs (Figure 1) separately as well as for the entire African continent. A total of 1754 validation samples were reserved and was only available to the validation team. Further, the areas computed for the 55 countries of Africa were compared with areas available from UN FAO.
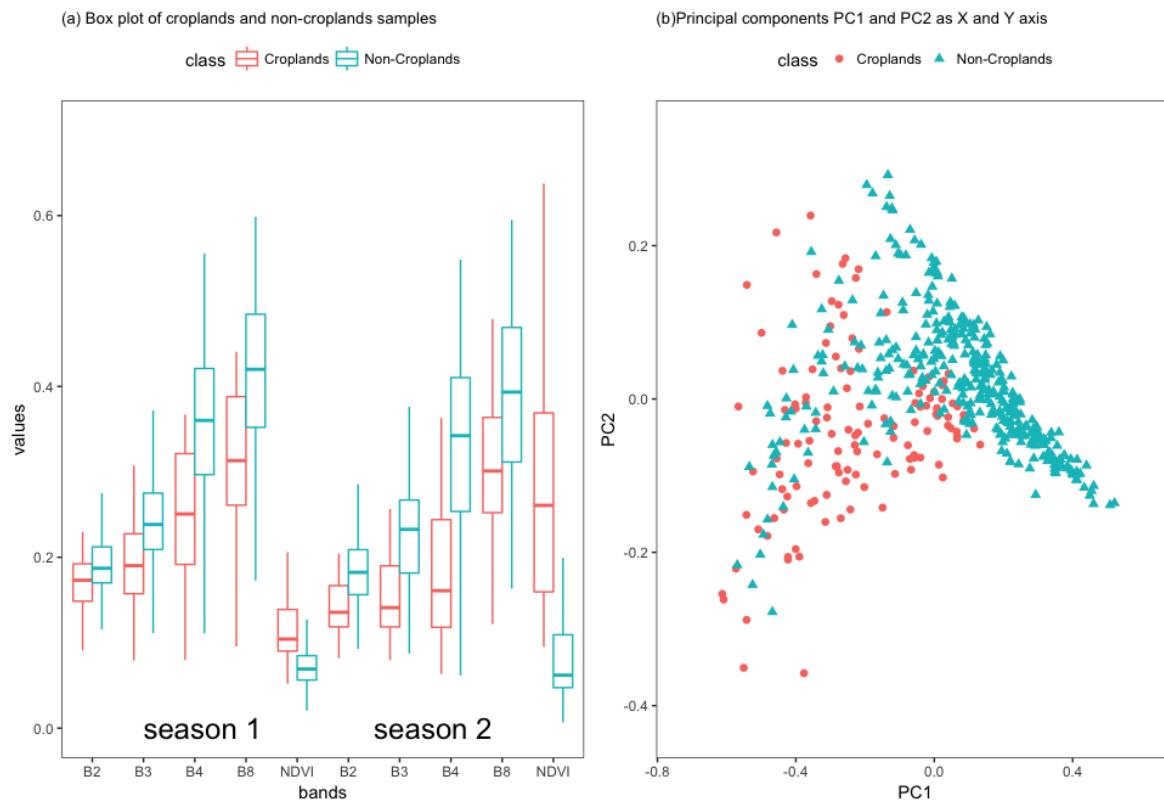
Figure 3. Creating knowledge-base to separate croplands from non-croplands involving: (**a**) waveband reflectivity and NDVI of the two seasons; and (**b**) principal component plot. The above knowledge-base is illustrated for a sub-area in one of the seven zones shown (Figure 1) of Africa.

3. Methods

There is no single classifier that is best applicable to cropland mapping [13,43] as a result of their various strengths and limitations [15], specifically for large areas [8,9]. As result, a combination of different classification techniques [47,76] were investigated. Earlier, several authors [49,50] have explored this combination of methods for land use/land cover classification over large areas, but not for cropland classification over large areas yet. In this section, an integrated methodology of pixel-based classifiers (Random Forest, Support Vector Machines) and object-based Recursive Hierarchical Segmentation is outlined in the flowchart (Figure 4) and described in the following sub-sections:

3.1. Overview of the Methodology

A comprehensive overview of the methodology is shown in Figure 4. As it is difficult to apply a single classifier over large areas (Figure 1) for cropland mapping, an ensemble of machine learning algorithms were investigated [47,76]. Specifically, the integration of pixel-based and the object-based classifiers for large area land cover mapping has been explored by several authors [49,50]. Both pixel-based and object-oriented classifiers require a large amount of reference training data (Section 2.3, Figure 1), which we established through several endeavors as discussed in Section 2 and its sub-sections.

1. 30-m mosaic (11 bands) was built using Sentinel-2 and Landsat-8 data (Section 2.2) for period 1 (January–June, 2016) and period 2 (July–December, 2015);
2. Random Forest and Support Vector Machines (Section 3.1) were used to classify input bands for croplands versus non-croplands;

3.   Using same bands as inputs, recursive hierarchical segmentation (Section 3.2) was carried out in 1° by 1° grid units on NASA pleiades supercomputer;

4.   The pixel-based classification was integrated with object-based segmentation into cropland extent map (Section 3.3) for further assessment (Section 3.4)

5.   We compared derived cropland areas with country-wise statistics from other sources in Section 3.5 and explored the consistency between GFSAD30AFCE map and other reference maps in Section 3.6.

6.   30-m Cropland extent product is released through the NASA Land Processes Distributed Active Archive Center (LP DAAC) at: https://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30AFCE. 001 and can be viewed at: https://croplands.org.
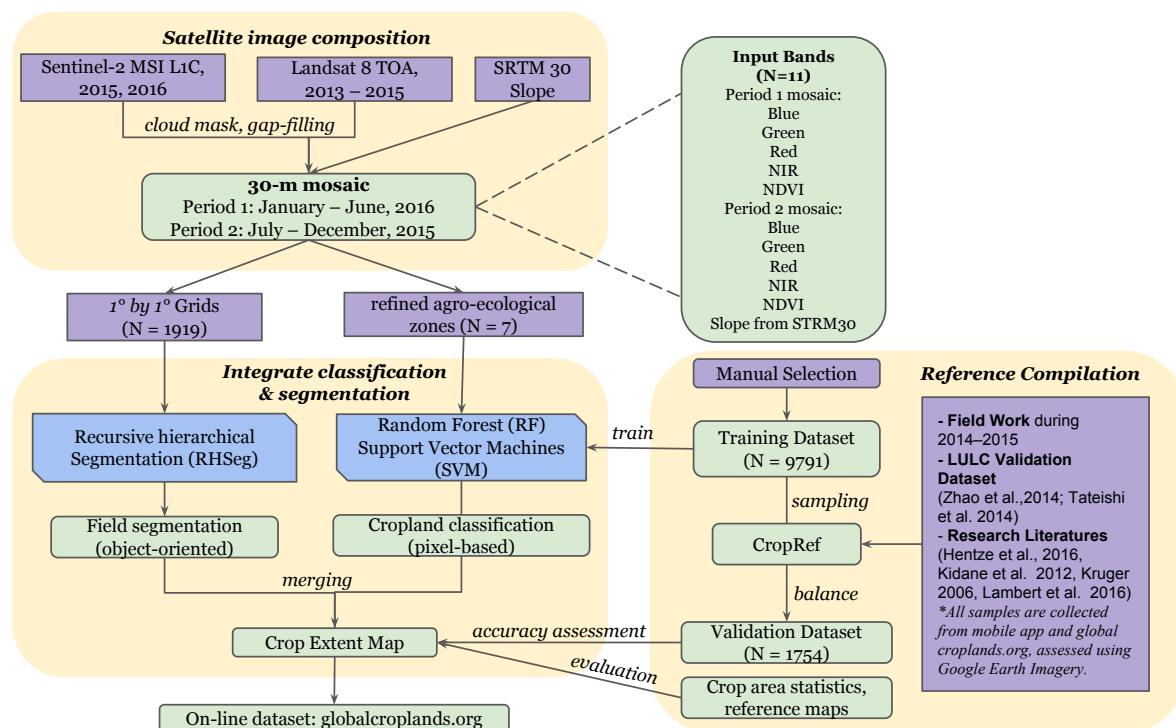

Figure 4. Overview of methodology for cropland extent mapping. The study integrates pixel-based classification involving the Random Forest (RF) and Support Vector Machines (SVM) with object-oriented Recursive Hierarchical Image Segmentation (RHSeg). The chart also shows the reference and training dataset used.

3.2. Pixel-Based Classifier: Random Forest (RF) and Support Vector Machine (SVM)

The Random Forest classifier uses bootstrap aggregating (bagging) to form an ensemble of trees by searching random subspaces from the given features and the best splitting of the nodes by minimizing the correlation between the trees. It is more robust, relatively faster in speed of classification, and easier to implement than many other classifiers [77]. Accurate land cover classification and better performance of the RF models have been described by many researchers [11,77–79].

RF was used to classify croplands in the 7 RAEZ's, as shown in Figure 1. Initially, 5511 training samples described in Section 2.3 were used for the first run of RF algorithm on GEE. In order to improve classification results, ~500-600 trees were used and the number of training samples were increased to varying degrees for each individual zone.

RF classified results were visually compared to other reference maps (GlobCover, GRIPC, GLC30, Google Earth Imagery). Based on these comparisons, the training polygon set was refined by creating additional croplands and non-croplands training polygons by drawing them in GEE editor.

After additional training polygons were added, the classification was run again. This iterative step is time-consuming, depending on the complexity of the landscape. For example, in the rainfed areas of central Africa like Tanzania, the rainfed cropland areas are mixed with natural vegetation and bare land. In such places, the training sample selection was repeated 4 times to achieve satisfactory results. For all 7 RAEZs, another 4280 polygons were added following this iterative procedure. Overall, we used 9791 training samples/polygons across the entire African continent.

Occasionally, overfitting issues happened in RF because input features are heavily correlated in specific areas; to correct for this a SVM classifier [80] with a Linear kernel was used in the problematic regions to replace RF results. SVM has also been reported to work significantly better with smaller, intelligently selected, training samples than RF in literature [81].

The pixel-based classifiers (RFs and SVMs) were run on GEE. Cloud computing offers the power of computing by linking 1000s of computers, allowing parallel processing and thus enabling the classification of individual zones with 30-m pixels in matter of hours.

3.3. Recursive Hierarchical Image Segmentation (RHSeg)

Section 3.2 discussed pixel-based classifiers which are fast and are scalable over large areas on cloud computing facilities like GEE. However, the pixel-based classification results inevitably include "salt and pepper" noise and dis-jointed farm fragments in practice. Object-based analysis can improve salt and pepper effects and increase classification accuracies over pixel-based image classification [82,83]. Image segmentation gathers several similar neighbor pixels together as objects, and categorizes or labels objects, which would be further labelled as croplands or non-croplands in the integration step with pixel-based classification in Section 3.2. Image segmentation procedures have many implementation [82,83] with very high memory and CPU requirements. GEE provides APIs related to image segmentation such as region grow [84]. However, image segmentation for entire Africa at 30-m is beyond GEE's existing capacity so we utilized NASA supercomputer facilities [85] to implement intensive segmentation over large areas.

In this study, Recursive Hierarchical Segmentation (RHSeg) software [86] was adopted to extract object information from 30-m input imagery. RHSeg is an approximation of the Hierarchical Segmentation (HSeg) algorithm that recursively subdivides large images into smaller subimages that can be effectively processed by HSeg. RHSeg then blends the results from the subimages to produce a hierarchical segmentation of the entire large image. HSeg utilizes an iterative region growing approach to produce hierarchical image segmentations. HSeg is unique in that it alternates merges of spatially adjacent and non-spatially adjacent regions to provide a simultaneous segmentation and classification. The addition of merging non-spatially adjacent regions helps stabilize the segmentation result by providing a larger sampling of each spectral class type. Other hierarchical classification strategies have been tested by several researchers with a series of per-class classifiers to minimize the effect of spectral confusion among different land cover classes [50,87].

The merging of spatially non-adjacent regions in RHSeg leads to heavy computational demands. In order to expand its capability from regional-size to continental scale, a grid scheme (Figure 1) was applied to subset the 30-m mosaic dataset created in Section 2.2, covering the non-desert areas of Africa into 1919 smaller pieces using GEE. Each scene was a 10-band image for entire Africa without the slope band (Figure 2, Table 1) at 30-m resolution, about 4000 columns by 4000 rows in size which was then used by RHSeg as input for segmentation. Generating the segmentation of these input datasets took about 74 hours using 64 CPUs on the Pleiades and Discover NASA Supercomputers under the parallel mode supported by RHSeg.

Noting that some image scenes had a large percentage of water pixels or pixels masked out due to clouds, we realized that more consistent and accurate results could be obtained by selecting results from the RHSeg segmentation hierarchy based on merge threshold instead of the number of regions. Based on the analysis of 12 representative 30-m foot-print images across Africa, we found

that a merge threshold of 15.0 selected the most suitable RHSeg segmentation hierarchy layer for agricultural application.

3.4. Integration of Pixel-Based Classification and Object-Based Segmentation

Every segment in the outputs of RHSeg at the selected hierarchical level consists a group of pixels with a unique id (region label), which will be further labeled as "cropland" or "non-cropland" patch when the segment was overlaid with pixel-based classification results. In order to merge pixel-based classification and field boundary information from segmentation, we reassigned the pixel values for individual segments according to the following rules which were developed based on trial and error in 12 images under different landscapes across the continent:

If > 85% of the pixels in a segmented patch are classified as 'cropland', the whole patch was assigned to 'cropland';

If < 15% of the pixels in a segmented patch are classified as 'cropland', the whole patch was assigned to 'non-cropland';

If either condition is not met, the pixel-based classification results will be unchanged in the final crop extent map, resulting in mixed cropland and non-croplands pixels in one patch.

The example shown in Figure 5 highlights the value of the merging steps above to produce the final cropland extent map. Pixel-based classification of croplands (green) covered most highly vegetated areas, however, some cropland-pixels were missing because of cropland heterogeneity and spectral contamination among neighboring pixels (Figure 5a). In Figure 5b, the RHSeg segmentation layer is better able to determine whether pixels belong to the same field (random coloring). The results from Figure 5a,b are merged to produce a more refined and complete boundaries of cropland fields (green) (Figure 5c) which has better consistency with the true color VHRI from Google Earth (Figure 5d).
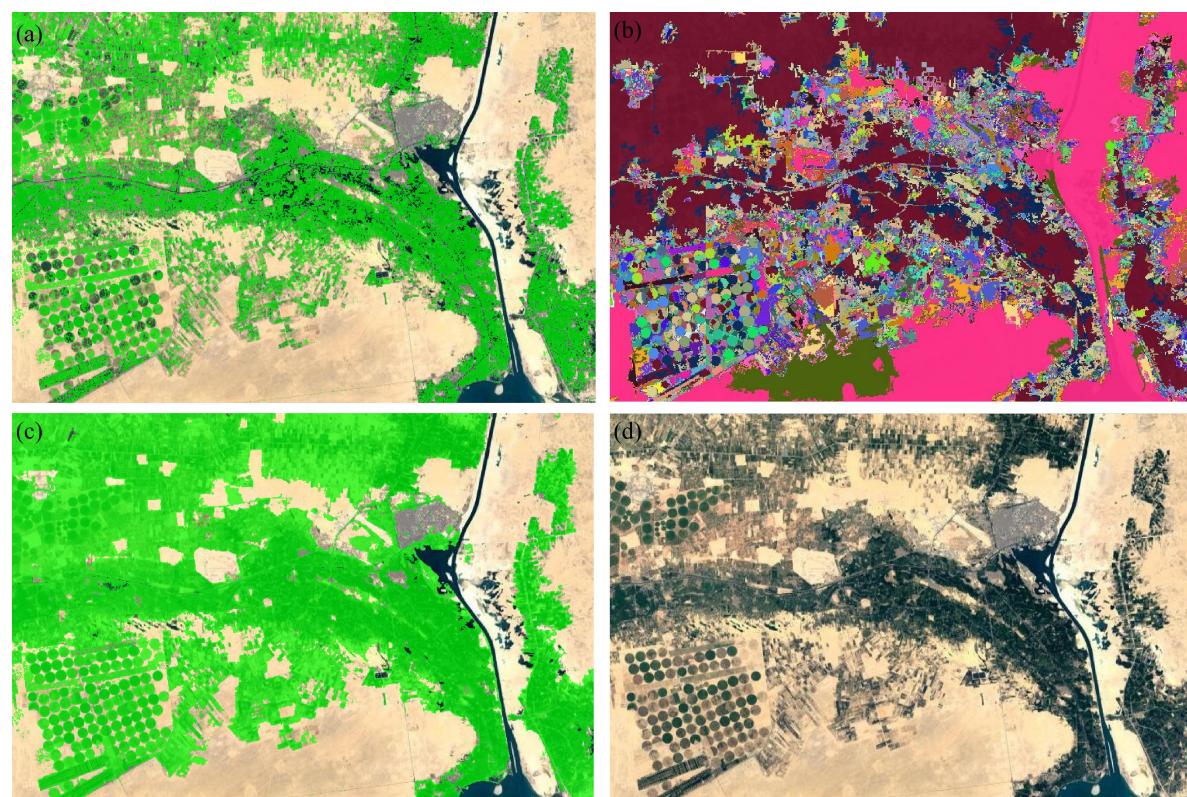

Figure 5. The example of (**a**) the pixel-based classification from random forest classifier; (**b**) the object-based RHSeg image segmentation result; (**c**) the merged results with RHSeg segmentation result with pixel-based Random Forest classification; and (**d**) a true color Google Earth Imagery for reference.

3.5. Accuracy Assessment

Map accuracy assessment is a key component of map production, especially when remote sensing data are utilized [88]. Validation exercises require high-quality reference validation data sets collected at appropriate spatial and temporal scales using random sample designs [89,90]. In addition to the accuracy analysis performed when evaluating the classification results to select the best algorithms and results, an independent validation of the product was performed. For the independent assessment, a total of 1754 samples were used to determine the accuracy of the final cropland extent map of Africa for all 7 RAEZ's (Figure 1). Error matrices were generated for each of the RAEZ's separately and also for the entire African continent providing producer's, user's, and overall accuracies. Further, the areas computed for the 55 countries of Africa were compared with areas available from UN FAO.

There are few basic important considerations that must be followed step by step in order to perform an assessment of the cropland thematic maps [89]. The process of validation usually starts with collection of a high-quality reference data independent of the training data that have already been used for mapping the same area. The reference samples have been collected from very high-resolution imagery or VHRI (sub-meter to 5-m) that were available for the entire continent through US National Geospatial Agency (NGA) though image interpretation that corresponded with the same year of mapping. It is better to adopt a continent specific sampling method to perform a meaningful assessment of global cropland products. For Africa, a stratified random sampling design [89] was used to distribute a balanced sample size using the following steps:

1. Stratified, random and balanced sampling: The African continent has been divided into 7 refined agro-ecological zones or RAEZs (Figure 1) for stratified random sampling. Due to a large crop diversity across RAEZ's (Figure 1) there is high variability in their growing periods and crop distribution. Therefore, to maintain balanced sampling for each zone, samples have been randomly distributed in each zone. The question of how many samples are sufficient to achieve statistically valid accuracy results is described in next point below.
2. Sample Size: The sample size has been chosen based on the analysis of incrementing minimum number of samples. Initially, first 50 samples were chosen as minimum number for all the 7 RAE's and then incremented in steps with another 50 more samples. A few RAEZ's in Africa have little cropland distribution so that 50 samples were enough to achieve a valid assessment. However, other RAEZ's needed up to 250 samples for their assessment. Beyond 250 samples, accuracies of all RAEZ's become asymptotic. Overall, for Africa, total 1754 samples were used from 7 RAEZ's.
3. Sample unit: The sample unit for a given validation sample must be a group of pixels (at least 3×3 pixels of 30-m resolution) in order to minimize the impact of positional accuracy [88]. This sampling unit is a 3×3 homogeneous window containing one class. If a sample at this step was recognized to be a mixed patch of cropland and non-cropland, it had to be excluded from the validation dataset in the accuracy assessment since heterogeneous windows were not considered, however excluding them is the best practical choice for accuracy assessment.
4. Sampling was balanced to keep the proportion of the cropland versus non-cropland samples close to the proportion of the cropland versus non-cropland area from the product layer to be validated.
5. Validation samples are created independently from training samples described in Section 2.3, by a different team.

The performance of the different approaches was assessed by two complementary criteria, namely the accuracy assessment and across-site robustness. Two different metrics, derived from the confusion matrix, were selected for the overall accuracy (OA) assessment. The OA evaluated the overall effectiveness of the algorithm, while the F-score measured the accuracy of a class using the precision

and recall measures. For each of the 7 RAEZ's (Figure 1) of Africa, the study establishes error matrices that provides user's (UA), producer's (PA), and overall accuracies (OA) as following equations:

$$OA = \frac{S_d}{n} \times 100\% \tag{1}$$

$$UA = \frac{X_{ij}}{X_j} \times 100\% \tag{2}$$

$$PA = \frac{X_{ij}}{X_i} \times 100\% \tag{3}$$

$$F_{score} = 2 \times \frac{UA \times PA}{UA + PA} \tag{4}$$

where S_d is the total number of correctly-classified pixels, n = total number of validation pixels, X_{ij} = observation in row i column j; X_i = marginal total of row i; X_j = marginal total of column j.

3.6. Calculation of Actual Cropland Areas and Comparison with Areas from Other Sources

Generating cropland areas such as at national and sub-national levels is of great importance in food security studies. In Google Earth Engine, we convert the crop extent map to a crop area map where the pixel value represents the actual crop area converting the map to Lambert Azimuthal (equal-area) projection. In order to derived the country level cropland area statistics from the 30-m crop extent map of Africa, we used the Global Administrative Unit Layers (GAUL) from UN FAO as country boundaries to create Table 4 as well as statistics from other sources, including AquaStat [91], Mirca2000 [92], GRIPC [16] and GLC30 [45].

3.7. Consistency between GFSAD30AFCE Product and Four Existing Crop Maps

The GFSAD30AF product was also compared with other LULC/Cropland products that were published recently to establish consistency between the products. First, we remapped four existing global land cover map products according to their individual classification schemes (Table 2):

- Global Land Cover Map for 2009 (GlobCover 2009) [39]. Class 11, 14 were reclassified as "croplands" and other land cover classes were reclassified as "non-croplands";
- Global rainfed, irrigated, and paddy croplands map GRIPC [16]. All agricultural classes include rainfed, irrigated and paddy were combined as "croplands" and other classes were "non-croplands";
- 30-m global land-cover map FROM-GLC [48]. Level 1 class 10 and Level 2 Bare-cropland 94 were combined as "croplands" and other classes were "non-croplands"; and
- Global land cover GLC30 [45]. Class 10 was combined as "croplands" and other classes were "non-croplands".

To unify spatial resolution of cropland maps, the Cropland Extent Map and these four cropland maps were all resampled to 30-m resolution for comparison. In addition to visual comparison illustrations, we also evaluated statistical agreements between these cropland maps. We generated 12,627 random points across the classification extent and sampled the cropland classes from five cropland extent maps to build similarity metrics.

Table 2. Remapped land cover classes of other cropland or land use/land cover (LULC) products used to compare with the GFSAD30AFCE product of this study.

Cropland/LULC Maps	Resolution	Code	Class Name
GlobCover 2009 v2.3	300 m	11	Post-flooding or irrigated croplands
		14	Rainfed croplands
		20	Mosaic cropland/vegetation
		30	Mosaic vegetation/croplands
GRIPC	500 m	1	Rain-fed croplands
		2	Irrigated croplands
		3	Paddy croplands
FROMGC Level 2	30 m	11	Rice
		12	Greenhouse
		13	Other Crop
		94	Bare-cropland
GLC30	30 m	1	Cultivated Land

4. Results

The study produced a nominal 30-m cropland extent product (Figure 6; croplands.org) of the entire African continent using Sentinel-2 and Landsat-8 data for the year 2015. In the following sub-sections, we will discuss this product, referred to as, the Global Food Security-support Analysis Data @ 30-m of Africa, Cropland Extent (GFSAD30AFCE; Figure 6) product, its accuracies, areas derived from it, and comparison of areas with areas reported through National and sub-National statistics as reported by the Food and Agricultural Organization (FAO) of the United Nations (UN). We will also compare the GFSAD30AFCE product with other cropland and/or land use/land cover (LULC) products where cropland classes were mapped.
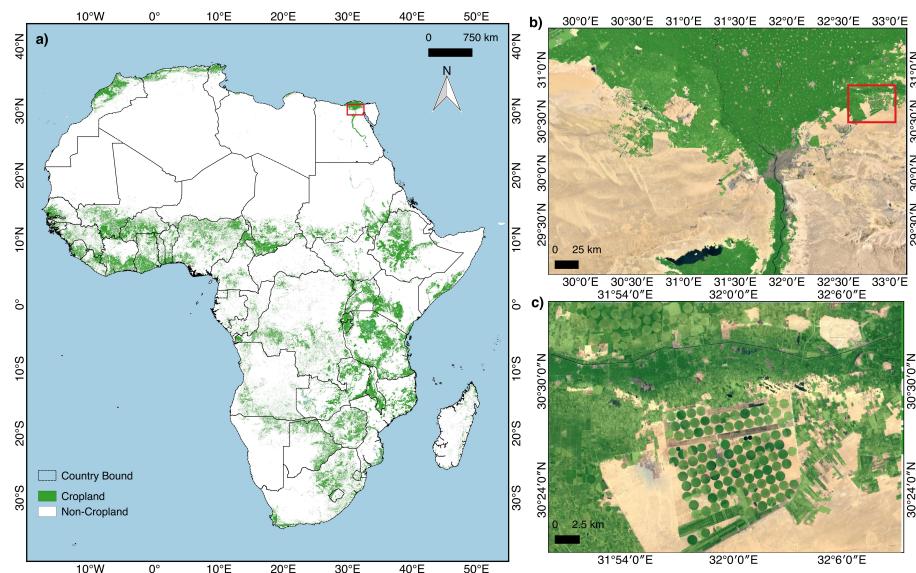

Figure 6. Global Food Security-support Analysis Data @ 30-m of Africa, Cropland Extent product (**a**). Full resolution of 30-m cropland extent can be visualized by zooming-in to specific areas as illustrated in right panel (**b**,**c**). For any area in Africa, croplands can be visualized by zooming into specific areas in croplands.org.

4.1. GFSAD30AFCE Product

The Global Food Security-support Analysis Data @ 30-m of Africa, Cropland Extent (GFSAD30AFCE; Figure 6), produced by combining the pixel-based (RF, SVM) and Object-based

segmentation algorithm (RHSeg), is accessible at https://croplands.org/app/map. The data will also be soon made available for download through NASA's Land Processes Distributed Active Archive Center (LP DAAC). The product year is referred to as nominal 2015 since most Sentinel-2 images used in processing were from July 2015 to June 2016. Data users can also browse the online version of the products at croplands.org.

On the African continent, croplands are primarily dominant throughout West Africa, along the great Lakes of Africa (Lake Victoria and Lake Tanganyika), South Africa, Southern Africa, along the coasts of North Africa, and all along the Nile Basin (Figure 6). The Sahara Desert, Kalahari Desert, and overwhelming proportion of the Congo rain forests have almost no croplands (Figure 6).

4.2. GFSAD30AFCE Product Accuracies

This final cropland extent product of Africa (GFSAD30AFCE) was systematically tested for accuracies (Table 3) by independent validation datasets in each of the 7 refined agro-ecological zones or RAEZs (Figure 1). For the entire African continent, the weighted overall accuracy was 94.5% with producer's accuracy of 85.9% (errors of omissions of 14.1%) and user's accuracy of 68.5% (errors of commissions of 31.5%) for the cropland class (Table 3). When considering all 7 RAEZs, the overall accuracies varied between 90.8% and 96.8%, Producer's accuracies varied between 60.7% and 94.9%, and user's accuracies varies between 53.3% and 89.6% for cropland class (Table 3). The F score ranged between 0.65 and 0.9.

Table 3. Independent Accuracy Assessment of GFSAD30 Cropland Extent product of Africa (GFSAD30AFCE).

Zone 1, % of TNCA* = 9.1%

		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%

		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%

		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%

		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%

		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%

		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%

		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%

		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: * TNCA (Total Net Croplands Area) = 313 Mha

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

Across RAEZs (Table 3), user's accuracies (commission errors) were significantly lower than producer's accuracies (omission errors). This was mainly because when training the random forest algorithm, we tweaked it to capture as much croplands as possible, thus ensuring high producer's accuracies (or low omission errors for the cropland class) across RAEZ's. In this process, the compromise was that some non-croplands were included as croplands, resulting in lower user's accuracies (or higher commission errors) for the cropland class. Ideally, an algorithm should optimize a classification to balance producer's and user's accuracies. However, the goal of this project is to map almost all croplands including fallow croplands. As a result, we aimed for high producer's accuracies (low errors of omissions) for the cropland class across zones (Table 3) and achieved it for most RAEZs, as evidenced by a continent-wide producer's accuracy of 85.9% (Table 3).

4.3. Cropland Areas and Comparison with Statistics from Other Sources

GFSAD30AFCE can accurately estimate cropland areas by nation or sub-national regions (e.g., state, district, county, village). Here we calculated the cropland areas by country in Africa (Table 4) for comparison with survey-based statistical area from UN FAO. Users can make use of this product to do their own computation of sub-national statistics anywhere in Africa at all kinds of administrative level and compare them to reliable reference data.

Table 4 shows a country-wise cropland area statistics of all 55 African countries generated from this study using GFSAD30AFCE product of year 2015 (Figure 6) and compared with the national census data based MIRCA2000 [92] which were updated in the year 2015 (Stefan Siebert and Portmann, personal communication). Overall, the entire African continent had total net cropland area (TNCA) of 313 million hectares (Mha). Five countries (Nigeria, Ethiopia, Sudan, Tanzania, and South Africa) constitute 40% of all cropland areas of Africa and each have 5% or more of total Africa's net cropland area (TNCA) of 313 Mha (Table 4). Nigeria is the leading cropland country in Africa with 11.4% of the 313 Mha (Table 4). Ethiopia is second with 8.21%. However, crop productivities will depend on numerous factors such as soils, whether they are irrigated or rainfed, management issues (e.g., inputs such a N, K, P), and also climate and plant genetics. Thereby, larger cropland area does not necessarily mean greater crop productivity. There are 12 countries (DR Congo, Mali, Zimbabwe, Kenya, Morocco, Algeria, Niger, Zambia, Uganda, Mozambique, Burkina Faso, Chad) which have above 2% but below 5% of Africa's TNCA of 313 Mha. The remaining 38 African countries have less than 2% of Africa's NCA. The overwhelming proportion (94%) of the cropland areas are in just 25 of the 55 countries (Table 4).

For 48 of the 55 countries (7 "outliers" countries removed) there was a strong relationship between the GFSAD30AFCE product produced cropland areas versus the MIRCA2000 produced cropland areas (Figure 7) with an R-square value of 0.78. When all 55 countries are considered, this relationship provides an R-square value of 0.65. The countries where GFSAD30AFCE under-estimated croplands include Cote d'Voire, Uganda, Cameroun, Ghana, and Tunisia (Figure 7). The countries where GFSAD30AFCE over-estimated croplands include Malawi, Kenya, Mozambique, and Egypt to mention few names (Figure 7). Causes of this variability are many. Besides uncertainties among the input data and methodology, the GFSAD30AFCE product and the national statistics differ due to these reasons:

- Different definition of "croplands" class: GFSAD30AFCE product as per definition, includes all agricultural annual standing croplands, cropland fallows, and permanent plantation crops whereas cropland areas reported in statistics may not include cropland fallows;
- Different time: GFSAD30AFCE incorporate the latest cultivated area in 2015–2016 as well as the croplands fallows whereas country reported cropland areas may happen in other years.

Table 4. Total net cropland areas (TNCA) of the African countries derived from the global food security-support analysis data @ 30-m cropland extent product (GFSAD30AFCE) and compared with other cropland area sources.

Country	Land (Mha) FAO-GAUL	Total Net Cropland Area (TNCA, Mha)					
		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**

Note: FAO GAUL = The Food and Agricultural Organization's The Global Administrative Unit Layers (GAUL); GFSAD30AFCE = global food security support analysis data @ 30-m (this study); GFSAD250 = global food security support-analysis data @ 250-m [15]; GRIPC = Global rain-fed, irrigated, and paddy croplands [16]; MIRCA2000 = Global data set of monthly irrigated and rainfed crop areas around 2000, revised for year 2015 in this study [92]; GLC30 = Global land cover mapping at 30m resolution [45].

There are a number of other reasons for discrepancies between remote sensing and non-remote sensing sources [1,8,13,23,43]. We suggest that a detailed investigation should be conducted on this aspect to see why uncertainties exist and how to overcome them. More detailed assessment of such variability is beyond the goal of this study. On average, the GFSAD30AFCE determined about 35% higher cropland areas relative to national statistics reported by Portmann et al., and UN FAO. It is important to note that the GFSAD30AFCE of this study provided TNCA of the continent as 313 Mha, which is 5.7% higher than our earlier MODIS 250-m data based estimate of 296 Mha [15]. Other studies reported far less cropland areas for Africa, which were (Table 4): 232 Mha (MIRCA), 211 Mha (FAO), 202 Mha (GRIPC), and 223 Mha (GLC30). Therefore, these estimates were lower by about 26% to 35% relative to GFSAD30AFCE product. Various factors may contribute to such discrepancies: 1. MIRCA and FAO UN statistics were derived from a combination of national reports and their synthesis using some remote sensing, GIS and field visits. MIRCA2000 is a derived gridded dataset based on the FAOSTAT database [92]. 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, and resource limitations, the data lacks objectivity in many countries, resulting 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 is determined by the number of steps required to walk around them [93]; 2. GRIPC [16] also maps croplands but using 500-m MODIS data and using different definition and methodologies other than GFSAD30AFCE; 3. GLC30 does include croplands class [45], and its focus is land use and land cover so a lot of uncultivated, low vegetated croplands were identified as shrubs or grass instead of croplands.
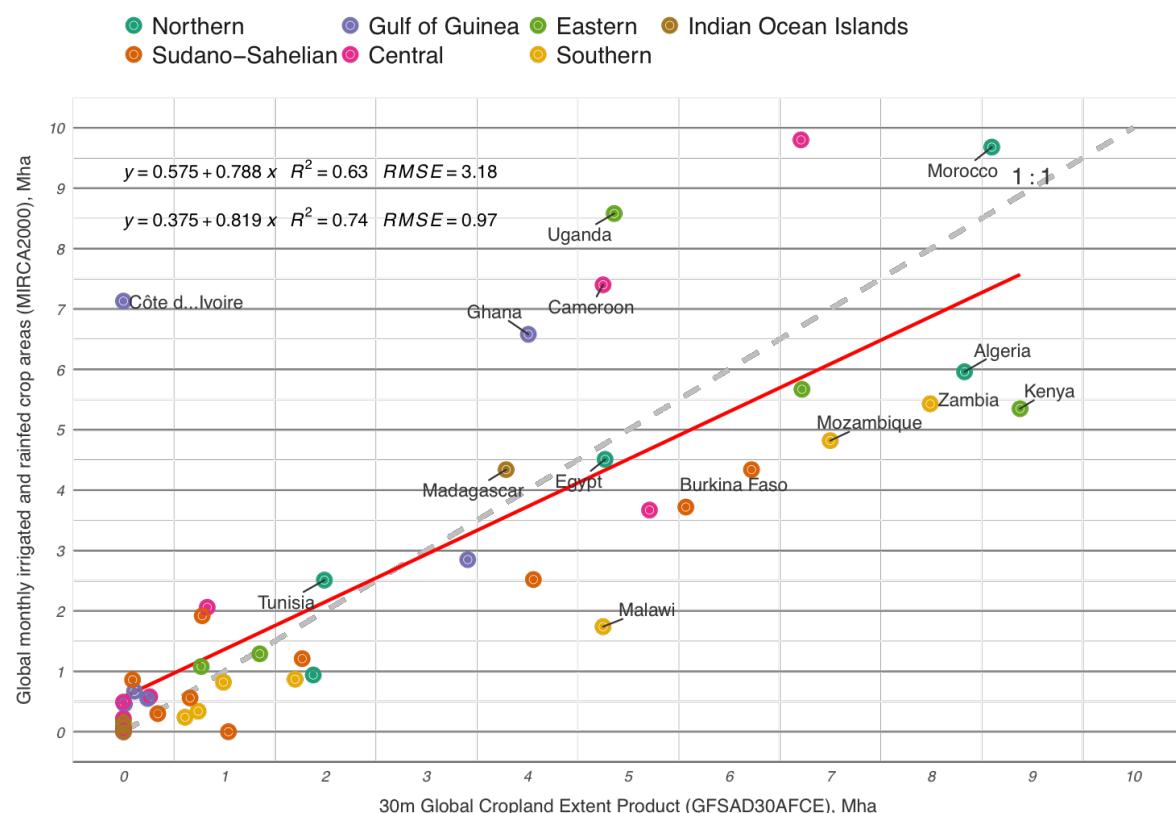

Figure 7. Scatter plot of GFSAD30AFCE derived cropland areas versus MIRCA2000 (Personal communication with Dr. Siebert and Dr. Portmann by Dr. Thenkabail) derived cropland areas country by country for Africa.

4.4. Consistency between GFSAD30AFCE Product and Four Existing Crop Maps

A similarity analysis was conducted where GFSAD30AFCE product was compared with each of the other four products (Table 2) using the 12,627 random samples spread across the African continent. The results showed that GFSAD30AFCE product matches with GLC30 product for 77.3% of the samples while matches with GRIPC500 product with 68.8% samples for the cropland class (Table 5). For the other two products (Globecover2009 and FROMGC), the cropland samples only matched 60% (Table 5). The discrepancies between the products include different reference year of the dataset, cropland definition and methodologies, resolution of the datasets, and a host of other factors; the differences between these products should be investigated further in the future.

Table 5. Similarity analysis comparing the GFSAD30AFCE product of this study with four other products using 12,627 random samples.

(a) Confusion Matrix Obtained with GlobCover 2009 (300-m, Arino et al., 2007)			
	Cropland	**Non-Cropland**	**User's Accuracy (%)**
Cropland	1125	585	65.79%
Non-Cropland	720	10,395	93.52%
Producer's Accuracy (%)	60.98%	94.67%	
(b) Confusion Matrix Obtained with GRIPC (500-m, Salmon et al., 2015)			
	Cropland	**Non-Cropland**	**User's Accuracy (%)**
Cropland	1267	443	74.09%
Non-Cropland	576	10,539	94.82%
Producer's Accuracy (%)	68.75%	95.97%	
(c) Confusion Matrix Obtained with FROMGC (30-m, Yu et al., 2013)			
	Cropland	**Non-Cropland**	**User's Accuracy (%)**
Cropland	1009	701	59.01%
Non-Cropland	676	10,439	93.92%
Producer's Accuracy (%)	59.88%	93.71%	
(d) Confusion Matrix Obtained with GLC30 (30-m, Chen et al., 2015)			
	Cropland	**Non-Cropland**	**User's Accuracy (%)**
Cropland	1409	301	82.40%
Non-Cropland	413	10,702	96.28%
Producer's Accuracy (%)	77.33%	97.26%	

Figure 8 shows a visual comparison of each product's mapping of cropland pixels (highlighted in green) using true color Google Earth Imagery as a background reference under three different landscapes: Egyptian irrigation area, South Africa irrigation area, Cote d'Ivoire mixed agricultural area. GFSADAFCE maps greatly outperformed the coarser resolution products which is significant considering these maps are frequently used in applications that monitor agricultural landscapes. Also, since images from the year 2015 were used in this study, it also covers newly cultivated farmlands which were not mapped in higher resolution product such as GLC30. Furthermore, the training/validation datasets of GFSAD will be also published with croplands maps, which means these training samples can be reused/expanded to update the latest cropland extent when necessary by using Google Earth Engine Cloud-based image composition and classification tools.
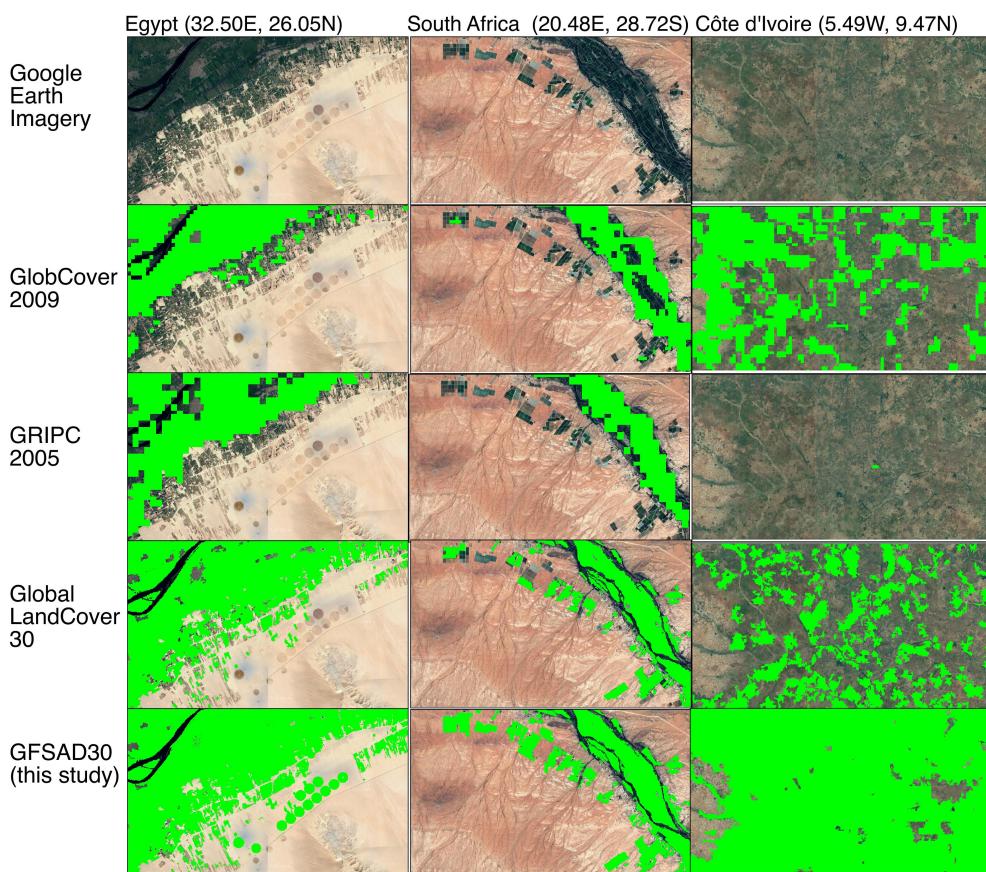
Figure 8. A visual comparison of all crop extent products (shown in green) overlaid on Google Earth Imagery.

5. Discussion

Although the value of this approach is evident, there are still problems in GFSAD30 cropland extent; some of these are discussed below. There were insufficient samples to reflect the diversity of croplands for certain regions and as a consequence, confusion between cropland and non-cropland classes exist. In Africa, the diversity of spectral properties for croplands is very high (e.g., cropland fallows in desert-margins of Sahara versus cropland fallows of the forest-margins of the rain forests). Even though we gathered a very large sample size for training and validation in this project, we still had difficulties verifying some areas. It was difficult to separate rainfed croplands from seasonal grasses in the Sahel and Northern Guinea Savanna because it is difficult to discriminate between them using VHRI imagery. In the context of phenology signatures, it is also easy to confuse croplands with bare lands and grasslands. For example, barelands, grasslands, and rainfed croplands in Sahel are very difficult to discern due to sparse vegetation of all three classes. We resolved such issues by utilizing acquisitions of quality samples from field visits, data (reference maps and ground data) from a few published articles on detailed studies in some small portions of the landscape and VHRI acquired during the exact growing seasons. Also, fallows of different ages (<1-year fallow to 5-year fallow) all have different signatures, especially in the rain forests where greater age corresponds to greater natural vegetation.

With 30-m resolution, satellite imagery is still limited for small fragmented fields in Africa, where field boundaries hardly exist and are often adjoin similar looking sparse grasslands or barrenlands. This is a specific problem throughout the Sahel and in certain places on the Northern Guinea Savannas of Africa. In such cases, RHSeg fails to identify boundaries of crop fields using 30-m imagery; however such issues can be improved by applying RHSeg to 10-m Sentinel-2 data in the future.

Another approach for improving classification accuracies is to use more refined segments as units for classification. A solution might be to integrate the FAO Farming systems map [2], which provides a finer stratification and takes into account both agro-ecological and climatic characteristics. In our earlier study using MODIS 250-m data, we did use FAO agro-ecological zones (AEZ's) for stratification [15]. However, the MODIS approach allowed monthly composites to be created, which was infeasible here when using 10 to 30-m data. By adopting more frequent (e.g., 15-day, monthly) periodic composites, we will be able to work with more detailed AEZ zones rather than the 7 RAEZ's zones as used in this study. We expect this will increase classification accuracies further.

6. Conclusions

This paper presents a practical methodology for cropland extent mapping at 30-m for the entire African continent on Google Earth Engine. Five-bands (blue, green, red, NIR, and NDVI) from 10-day time series Sentinel-2 and 16-day time-series Landsat-8 were time-composited over each of the two crop growing periods (period 1: January–June 2016; period 2: July–December 2015) along with 30-m SRTM DEM data, resulting in a 11-band stack over entire Africa. This input data was then classified using two pixel-based supervised classifiers: Random Forests (RFs) and Support Vector Machines (SVMs) which were merged with RHSeg, an object-based segmentation algorithm. A total of 9791 training samples/polygons were used to train the supervised classifiers. A total of 1754 validation samples were used for assessing accuracies, errors, and uncertainties.

The study produced the first cropland extent map of Africa at nominal 30-m resolution for the nominal year 2015. The product is referred to as the Global Food Security-support Analysis Data@ 30-m of Africa, Cropland Extent (GFSAD30AFCE; Figure 6). The weighted overall accuracy of the entire Africa continent was 94.5% with producer's accuracy of 85.9% (errors of omissions of 14.1%) and user's accuracy of 68.5% (errors of commissions of 31.5%). Across the 7 zones of Africa, accuracies vary with: 90.8% –96.8% overall accuracies, 60.7% –94.9% producer's accuracies, and 53.3% –89.6% user's accuracies. The F-score ranged between 0.65 and 0.90 across all 7 zones.

Derived from GFSAD30AFCE, total net croplands areas (TNCA's) was 313 million hectares for the African continent for the year 2015. In comparison to other cropland products of the past for African continent these area estimates were 26% to 35% higher. Five countries constitute 40% of all cropland areas of Africa, Nigeria (11.4%), Ethiopia (8.2%), Sudan (7.3%), Tanzania (7.2%) and South Africa (6.4%). There are 12 countries (DR Congo, Mali, Zimbabwe, Kenya, Morocco, Algeria, Niger, Zambia, Uganda, Mozambique, Burkina Faso, Chad) which each have above 2% but below 5% of Africa's TNCA. The remaining 38 African countries each have less than 2% of Africa's TNCA. The GFSAD30AFCE cropland areas explained 65% –78% percent of variability in UN FAO country-wise cropland areas.

The GFSAD30AFCE products are viewable at: https://croplands.org/app/map. The GFSAD30AFCE (https://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30AFCE.001) is released through NASA's Land Processes Distributed Active Archive Center (LP DAAC) for download by user community.

Cloud-based computing platforms such as Google Earth Engine and new earth-observing satellites like the Sentinel-2 constellation have brought significant paradigm-shifts in LULC mapping and agricultural cropland mapping and monitoring. The production of standard static maps will be replaced by the dynamic creation of maps from big data using crowd sourced training samples, and cloud computing which will better serve land managers, NGO's and the scientific community.

Abbreviations

The following abbreviations are used in this manuscript:

GFSAD30: Global Food Security Analysis-Support Data Project
GEE: Google Earth Engine

References

1. Thenkabail, P.S.; Hanjra, M.A.; Dheeravath, V.; Gumma, M. 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 Sens.* **2010**, *2*, 211–261.

2. Fritz, S.; See, L.; Rembold, F. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *Int. J. Remote Sens.* **2010**, *31*, 2237–2256.

3. Herold, M.; Mayaux, P.; Woodcock, C.E.; Baccini, A.; Schmullius, C. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens. Environ.* **2008**, *112*, 2538–2556.

4. See, L.; Fritz, S.; You, L.; Ramankutty, N.; Herrero, M.; Justice, C.; Becker-Reshef, I.; Thornton, P.; Erb, K.; Gong, P.; et al. Improved global cropland data as an essential ingredient for food security. *Glob. Food Secur.* **2015**, *4*, 37–45.

5. Delrue, J.; Bydekerke, L.; Eerens, H.; Gilliams, S.; Piccard, I.; Swinnen, E. Crop mapping in countries with small-scale farming: A case study for West Shewa, Ethiopia. *Int. J. Remote Sens.* **2012**, *34*, 2566–2582.

6. Hannerz, F.; Lotsch, A. Assessment of land use and cropland inventories for Africa. In *CEEPA Discussion Papers*; University of Pretoria: Pretoria, South Africa, 2006; Volume 22.

7. Gallego, F.J.; Kussul, N.; Skakun, S.; Kravchenko, O.; Shelestov, A.; Kussul, O. Efficiency assessment of using satellite data for crop area estimation in Ukraine. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *29*, 22–30.

8. Thenkabail, P.S.; Wu, Z. An Automated Cropland Classification Algorithm (ACCA) for Tajikistan by Combining Landsat, MODIS, and Secondary Data. *Remote Sens.* **2012**, *4*, 2890–2918.

9. Wu, W.; Shibasaki, R.; Yang, P.; Zhou, Q.; Tang, H. Remotely sensed estimation of cropland in China: A comparison of the maps derived from four global land cover datasets. *Can. J. Remote Sens.* **2014**, *34*, 467–479.

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

11. Tian, S.; Zhang, X.; Tian, J.; Sun, Q. Random Forest Classification of Wetland Landcovers from Multi-Sensor Data in the Arid Region of Xinjiang, China. *Remote Sens.* **2016**, *8*, 954, doi:10.3390/rs8110954.

12. Pittman, K.; Hansen, M.C.; Becker-Reshef, I.; Potapov, P.V.; Justice, C.O. Estimating Global Cropland Extent with Multi-year MODIS Data. *Remote Sens.* **2010**, *2*, 1844–1863.

13. Thenkabail, P.S.; Biradar, C.M.; Noojipady, P.; Dheeravath, V.; Li, Y.; Velpuri, M.; Gumma, M.; Gangalakunta, O.R.P.; Turral, H.; Cai, X.; et al. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. *Int. J. Remote Sens.* **2009**, *30*, 3679–3733.

14. Teluguntla, P.; Thenkabail, P.S.; Xiong, J.; Gumma, M.K.; Congalton, R.G.; Oliphant, A.; Poehnelt, J.; Yadav, K.; Rao, M.N.; Massey, R. Spectral Matching Techniques (SMTs) and Automated Cropland Classification Algorithms (ACCAs) for Mapping Croplands of Australia using MODIS 250-m Time-series (2000–2015) Data. *Int. J. Digit. Earth* **2017**, 944–977.

15. Xiong, J.; Thenkabail, P.S.; Gumma, M.K.; Teluguntla, P.; Poehnelt, J.; Congalton, R.G.; Yadav, K.; Thau, D. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 225–244.

16. Salmon, J.M.; Friedl, M.A.; Frolking, S.; Wisser, D.; Douglas, E.M. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *38*, 321–334.

17. Alcántara, C.; Kuemmerle, T.; Prishchepov, A.V.; Radeloff, V.C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. *Remote Sens. Environ.* **2012**, *124*, 334–347.

18. Estel, S.; Kuemmerle, T.; Alcántara, C.; Levers, C.; Prishchepov, A.; Hostert, P. Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. *Remote Sens. Environ.* **2015**, *163*, 312–325.

19. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182.

20. Boryan, C.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US agriculture: The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto Int.* **2011**, *26*, 341–358.

21. Vintrou, E.; Desbrosse, A.; Bégué, A.; Traoré, S. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *14*, 83–93.

22. Dheeravath, V.; Thenkabail, P.S.; Thenkabail, P.S.; Noojipady, P.; Chandrakantha, G.; Reddy, G.P.O.; Gumma, M.K.; Biradar, C.M.; Velpuri, M.; Gumma, M.K. Irrigated areas of India derived using MODIS 500 m time series for the years 2001–2003. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, doi:10.1016/j.isprsjprs.2009.08.004.

23. Biradar, C.M.; Thenkabail, P.S.; Noojipady, P.; Li, Y.; Dheeravath, V.; Turral, H.; Velpuri, M.; Gumma, M.K.; Gangalakunta, O.R.P.; Cai, X.L.; et al. A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 114–129.

24. Lambert, M.J.; Waldner, F.; Defourny, P. Cropland Mapping over Sahelian and Sudanian Agrosystems: A Knowledge-Based Approach Using PROBA-V Time Series at 100-m. *Remote Sens.* **2016**, *8*, doi:10.3390/rs8030232.

25. Shao, Y.; Lunetta, R.S. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS J. Photogramm. Remote Sens.* **2012**, *70*, 78–87.

26. Kussul, N.; Skakun, S.; Shelestov, A. Regional scale crop mapping using multi-temporal satellite imagery. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 45–52.

27. Skakun, S.; Kussul, N.; Shelestov, A.Y.; Lavreniuk, M.; Kussul, O. Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *9*, 3712–3719.

28. Huang, C.; Davis, L.S.; Townshend, J.R.G. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2010**, *23*, 725–749.

29. Vintrou, E.; Ienco, D.; Bégué, A. Data mining, a promising tool for large-area cropland mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2132–2138.

30. Pan, Y.; Hu, T.; Zhu, X.; Zhang, J. Mapping cropland distributions using a hard and soft classification model. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4301–4312.

31. Van Niel, T.G.; McVicar, T.R. Determining temporal windows for crop discrimination with remote sensing: A case study in south-eastern Australia. *Comput. Electron. Agric.* **2004**, *45*, 91–108.

32. Conrad, C.; Dech, S.; Dubovyk, O.; Fritsch, S.; Klein, D.; Löw, F.; Schorcht, G.; Zeidler, J. Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multitemporal RapidEye images. *Comput. Electron. Agric.* **2014**, *103*, 63–74.

33. Löw, F.; Michel, U.; Dech, S.; Conrad, C. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines. *ISPRS J. Photogramm. Remote Sens.* **2013**, *85*, 102–119.

34. Matton, N.; Canto, G.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series. *Remote Sens.* **2015**, *7*, 13208–13232.

35. Peña-Barragán, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316.

36. Johnson, D.M.; Mueller, R. The 2009 Cropland Data Layer. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 1201–1205.

37. Kalensky, Z.D. AFRICOVER Land Cover Database and Map of Africa. *Can. J. Remote Sens.* **1998**, *24*, 292–297.

38. Bartholomé, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977.

39. Arino, O.; Gross, D.; Ranera, F.; Leroy, M.; Bicheron, P.; Brockman, C.; Defourny, P.; Vancutsem, C.; Achard, F.; Durieux, L.; et al. GlobCover: ESA service for global land cover from MERIS. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–27 July 2007; pp. 2412–2415.

40. Latham, J.; Cumani, R.; Rosati, I.; Bloise, M. *Global Land Cover Share (GLC-SHARE) Database Beta-Release Version 1.0-2014*; FAO: Rome, Italy, 2014. Available online: http://csdms.colorado.edu/wiki/Data:GLC-SHARE (accessed on 31 May 2017).

41. Friedl, M.A.; McIver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; et al. Global land cover mapping from MODIS: algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302.

42. Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps. *Remote Sens.* **2015**, *7*, 7959–7986.

43. Teluguntla, P.; Thenkabail, P.; Xiong, J.; Gumma, M.K.; Giri, C.; Milesi, C.; Ozdogan, M.; Congalton, R.; Yadav, K. CHAPTER 6—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. In *Remote Sensing Handbook (Volume II): Land Resources Monitoring, Modeling, and Mapping with Remote Sensing*; Thenkabail, P.S., Ed.; CRC Press: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2015; pp. 131–160.

44. Waldner, F.; De Abelleyra, D.; Verón, S.R.; Zhang, M.; Wu, B.; Plotnikov, D.; Bartalev, S.; Lavreniuk, M.; Skakun, S.; Kussul, N.; et al. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity. *Int. J. Remote Sens.* **2016**, *37*, 3196–3231.

45. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Zhang, W.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27.

46. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.; et al. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654.

47. Hansen, M.C.; Loveland, T.R. A review of large area monitoring of land cover change using Landsat data. *Remote Sens. Environ.* **2012**, *122*, 66–74.

48. Yu, L.; Wang, J.; Clinton, N.; Xin, Q.; Chen, Y.; Zhong, L.; Gong, P. FROM-GC: 30 m global cropland extent derived through multisource data integration. *Int. J. Digit. Earth* **2013**, *6*, 521–533, doi:10.1080/17538947.2013.822574.

49. Costa, H.; Carrão, H.; Bação, F.; Caetano, M. Combining per-pixel and object-based classifications for mapping land cover over large areas. *Int. J. Remote Sens.* **2014**, *35*, 738–753.

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

51. Malinverni, E.S.; Tassetti, A.N.; Mancini, A.; Zingaretti, P.; Frontoni, E.; Bernardini, A. Hybrid object-based approach for land use/land cover mapping using high spatial resolution imagery. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1025–1043.

52. Dingle Robertson, L.; King, D.J. Comparison of pixel- and object-based classification in land cover change mapping. *Int. J. Remote Sens.* **2011**, *32*, 1505–1529.

53. Ok, A.O.; Akar, O.; Gungor, O. Evaluation of random forest method for agricultural crop classification. *Eur. J. Remote Sens.* **2012**, *45*, doi:10.5721/EuJRS20124535.

54. De Wit, A.J.W.; Clevers, J.G.P.W. Efficiency and accuracy of per-field classification for operational crop mapping. *Int. J. Remote Sens.* **2010**, *25*, 4091–4112.

55. Castillejo-González, I.L.; López-Granados, F.; García-Ferrer, A.; Peña-Barragán, J.M.; Jurado-Expósito, M.; de la Orden, M.S.; González-Audicana, M. Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Comput. Electron. Agric.* **2009**, *68*, 207–215.

56. Marshall, M.T.; Husak, G.J.; Michaelsen, J.; Funk, C.; Pedreros, D.; Adoum, A. Testing a high-resolution satellite interpretation technique for crop area monitoring in developing countries. *Int. J. Remote Sens.* **2011**, *32*, 7997–8012.

57. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, doi:10.1016/j.rse.2017.06.031.

58. Lupien, J.R. *Agriculture Food and Nutrition for Africa–A Resource Book for Teachers of Agriculture*; FAO: Rome, Italy, 1997. Available online: http://www.fao.org/docrep/w0078e/w0078e00.htm (accessed on 12 October 2016).

59. Gerland, P.; Raftery, A.E.; ikova, H.E.; Li, N.; Gu, D.; Spoorenberg, T.; Alkema, L.; Fosdick, B.K.; Chunn, J.; Lalic, N.; et al. World population stabilization unlikely this century. *Science* **2014**, *346*, 234–237.

60. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36.

61. Immitzer, M.; Vuolo, F.; Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* **2016**, *8*, 166, doi:10.3390/rs8030166.

62. Battude, M.; Al Bitar, A.; Morin, D.; Cros, J.; Huc, M.; Sicre, C.M.; Le Dantec, V.; Demarez, V. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens. Environ.* **2016**, *184*, 668–681.

63. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.* **2015**, *7*, 12356–12379.

64. Valero, S.; Morin, D.; Inglada, J.; Sepulcre, G.; Arias, M.; Hagolle, O.; Dedieu, G.; Bontemps, S.; Defourny, P.; Koetz, B. Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions. *Remote Sens.* **2016**, *8*, 55, doi:10.3390/rs8010055.

65. Lohou, F.; Kergoat, L.; Guichard, F.; Boone, A.; Cappelaere, B.; Cohard, J.M.; Demarty, J.; Galle, S.; Grippa, M.; Peugeot, C.; et al. Surface response to rain events throughout the West African monsoon. *Atmos. Chem. Phys.* **2014**, *14*, 3883–3898.

66. Hentze, K.; Thonfeld, F.; Menz, G. Evaluating Crop Area Mapping from MODIS Time-Series as an Assessment Tool for Zimbabwe's "Fast Track Land Reform Programme". *PLoS ONE* **2016**, *11*, doi:10.1371/journal.pone.0156630.

67. Kidane, Y.; Stahlmann, R.; Beierkuhnlein, C. Vegetation dynamics, and land use and land cover change in the Bale Mountains, Ethiopia. *Environ. Monit. Assess.* **2012**, *184*, 7473–7489.

68. Kruger, A.C. Observed trends in daily precipitation indices in South Africa: 1910–2004. *Int. J. Climatol.* **2006**, *26*, 2275–2285.

69. 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. Precipitation Patterns in West Africa. *Mon. Weather Rev.* **1980**, *108*, 1567–1578.

70. D'Odorico, P.; Gonsamo, A.; Damm, A.; Schaepman, M.E. Experimental Evaluation of Sentinel-2 Spectral Response Functions for NDVI Time-Series Continuity. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1336–1348.

71. Van der Werff, H.; van der Meer, F. Sentinel-2A MSI and Landsat 8 OLI Provide Data Continuity for Geological Remote Sensing. *Remote Sens.* **2016**, *8*, 883, doi:10.3390/rs8110883.

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

73. Languille, F.; Déchoz, C.; Gaudel, A.; Greslou, D.; de Lussy, F.; Trémas, T.; Poulain, V. Sentinel-2 geometric image quality commissioning: First results. *Proc. SPIE* **2015**, *9643*, 964306, doi:10.1117/12.2194339.

74. Barazzetti, L.; Cuca, B.; Previtali, M. Evaluation of registration accuracy between Sentinel-2 and Landsat 8. *Proc. SPIE* **2016**, doi:10.1117/12.2241765.

75. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, doi:10.1029/2005RG000183.

76. Aitkenhead, M.J.; Aalders, I.H. Automating land cover mapping of Scotland using expert system and knowledge integration methods. *Remote Sens. Environ.* **2011**, *115*, 1285–1295.

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

78. Sharma, R.; Tateishi, R.; Hara, K.; Iizuka, K. Production of the Japan 30-m Land Cover Map of 2013–2015 Using a Random Forests-Based Feature Optimization Approach. *Remote Sens.* **2016**, *8*, 429, doi:10.3390/rs8050429.

79. Wessels, K.; van den Bergh, F.; Roy, D.; Salmon, B.; Steenkamp, K.; MacAlister, B.; Swanepoel, D.; Jewitt, D. Rapid Land Cover Map Updates Using Change Detection and Robust Random Forest Classifiers. *Remote Sens.* **2016**, *8*, 888.

80. Vapnik, V.N.; Vapnik, V. *Statistical Learning Theory*; Wiley: New York, NY, USA, 1998.

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

82. Im, J.; Jensen, J.R.; Tullis, J.A. Object-based change detection using correlation image analysis and image segmentation. *Int. J. Remote Sens.* **2008**, *29*, 399–423.

83. Stow, D.; Hamada, Y.; Coulter, L.; Anguelova, Z. Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. *Remote Sens. Environ.* **2008**, *112*, 1051–1061.

84. Espindola, G.; Câmara, G.; Reis, I.; Bins, L.; Monteiro, A. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *Int. J. Remote Sens.* **2006**, *27*, 3035–3040.

85. Nemani, R.; Votava, P.; Michaelis, A.; Melton, F.; Milesi, C. Collaborative supercomputing for global change science. *Eos Trans. Am. Geophys. Union* **2011**, *92*, 109–110.

86. Tilton, J.C.; Tarabalka, Y.; Montesano, P.M.; Gofman, E. Best Merge Region-Growing Segmentation with Integrated Nonadjacent Region Object Aggregation. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4454–4467.

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

88. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practice*, 2nd ed.; CRC/Taylor & Francis: Boca Raton, FL, USA, 2009; p. 183.

89. Congalton, R.G. Assessing Positional and Thematic Accuracies of Maps Generated from Remotely Sensed Data. In *"Remote Sensing Handbook" (Volume I): Remotely Sensed Data Characterization, Classification, and Accuracies*; Thenkabail, P.S., Ed.; CRC Press: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2015; pp. 583–602.

90. Thenkabail, P.S.; Knox, J.W.; Ozdogan, M.; Gumma, M.K.; Congalton, R.G.; Wu, Z.; Milesi, C.; Finkral, A.; Marshall, M.; Mariotto, I.; et al. Assessing Future Risks to Agricultural Productivity, Water Resources and Food Security: How Can Remote Sensing Help? *Photogramm. Eng. Remote Sens.* **2012**, *78*, 773–782.

91. Chapagain, A.K.; Hoekstra, A.Y. The global component of freshwater demand and supply: An assessment of virtual water flows between nations as a result of trade in agricultural and industrial products. *Water Int.* **2008**, *33*, 19–32.

92. Portmann, F.T.; Siebert, S.; Döll, P. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Glob. Biogeochem. Cycles* **2010**, *24*, 1–24.

93. Dorward, A.; Chirwa, E. A Review of Methods for Estimating Yield and Production Impacts. 2010. Available online: http://eprints.soas.ac.uk/16731/1/FISP%20Production%20Methodologies%20review%20Dec%20Final.pdf (accessed on 10 August 2016).