Constructing a 30m African Cropland Layer for 2016 by Integrating Multiple Remote sensing, crowdsourced, and Auxiliary Datasets

Mohsen Nabil a,b,c, Miao Zhang a, Bingfang Wu a,b, Jose Bofana a,b,d and Abdelrazek Elnashar a,b,e

aState Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences (AIRCAS), Beijing, China; bCollege of Resources and Environment, University of Chinese Academy of Sciences (UCAS), Beijing, China; cDivision of Agriculture Applications, Soils, and Marine (AASMD), National Authority for Remote Sensing & Space Sciences (NARSS), Cairo, Egypt; dFaculty of Agricultural Sciences, Catholic University of Mozambique, Cuamba, Mozambique; eDepartment of Natural Resources, Faculty of African Postgraduate Studies, Cairo University, Giza, Egypt

ABSTRACT
Despite its essential importance to various spatial agriculture and environmental applications, the information on actual cropland area and its geographical distribution remain highly uncertain over Africa among remote-sensing products. Each of the African regions has its unique physical and environmental limiting factors to accurate cropland mapping, which leads to high spatial discrepancies among remote sensing cropland products. Since no dataset could cope with all limitations, multiple datasets initially derived from various remote sensing sensors and classification techniques must be integrated into a more accurate cropland product than individual layers. Here, in the current study, four cropland products, produced initially from multiple sensors (e.g. Landsat-8 OLI, Sentinel-2 MSI, and PROBA–V) to cover the period (2015–2017), were integrated based on their cropland mapping accuracy to build a more accurate cropland layer. The four cropland layers’ accuracy was assessed at Agro-ecological zones units via an intensive reference dataset (17,592 samples). The most accurate cropland layer was then identified for each zone to construct the final cropland mask at 30 m resolution for the nominal year of 2016 over Africa. As a result, the new layer was produced in higher cropland mapping accuracy (overall accuracy = 91.64% and cropland’s F-score = 0.75). The layer mapped the African cropland area as 282 Mha (9.38% of the Continent area). Compared to earlier cropland synergy layers, the constructed cropland mask showed a considerable improvement in its spatial resolution (30 m instead of 250 m), mapping quality, and closeness to official statistics ($R^2 = 0.853$ and $RMSE = 2.85$ Mha). The final layer can be downloaded as described under the “Data Availability Statement” section.

CONTACT Bingfang Wu, wubf@radi.ac.cn State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences (AIRCAS), Beijing 100101, China

ARTICLE HISTORY
Received 12 January 2021
Accepted 1 April 2021

KEYWORDS
Cropland mapping: synergy mapping; land cover; accuracy assessment; multi-source data fusion

© 2021 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the CASEarth Strategic Priority Research Programme.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
1. Introduction

The accurate cropland mask is the basic layer in spatial agriculture and environmental monitoring projects (Fritz et al., 2013; Lu et al., 2020; Waldner et al., 2016). The final accuracy of mapping the crop types (Frolking et al., 2002), plant water requirement (Ozdogan, Yang, Allez, & Cervantes, 2010), crop intensification (Hao, Tang, Chen, Yu, & Wu, 2019), fertilization management, pesticides applications (Mutanga, Dube, & Galal, 2017), and yield estimation (Yu et al., 2020) is highly dependent on the accuracy of the cropland binary layer/mask that shows where the croplands are geographically distributed. This layer is also essential for tracking dynamic changes of croplands and analyzing their driving forces (Xu et al., 2018). The projects that target croplands spatial monitoring are of paramount importance to sustainable management and agriculture development, particularly in food crisis regions such as Africa, where half of the number of acutely food-insecure people live (FSIN, 2020).

The primary source of cropland maps is from remote sensing datasets; however, these datasets tend to disagree with each other on the actual cropland area and geographical distribution, particularly over Africa (Pérez-Hoyos, Rembold, Kerdiles, & Gallego, 2017). Furthermore, their accuracy of mapping cropland in Africa does not exceed 70% in best cases (Nabil et al., 2020; Xu et al., 2019). The factors that limit the accurate mapping of African cropland from remote sensing were investigated by several studies (Nabil et al., 2020; Wei, Lu, Wu, & Ru, 2020). The main limiting factors were the high heterogeneity of the African landscape, intensive clouds over some African regions (e.g. west African coast), the fragmented field size, and complex topography (Nabil et al., 2020). To overcome those limitations, multiple sensors and classification techniques need to be tested for each region to select the classification approach that yields the highest accuracy (Bofana et al., 2020). Overall, no single classification method could yield a high mapping accuracy for all agro-environmental conditions (Nabil et al., 2020; Xiong et al., 2017). Hence, integrating multiple remote sensing datasets into a more accurate hybrid/synergic map was proposed to improve cropland mapping accuracy (Chen et al., 2019; Clinton, Le, & Peng, 2015; Lu et al., 2017; Tsendbazar, De Bruin, Fritz, & Herold, 2015). The newly constructed map is assumed to overcome individual layers’ limitations and be more spatially accurate (Lu et al., 2020).

Several datasets’ integration methods were developed and can be summarized into two main categories: prediction methods and ranking methods. The first approach uses various prediction (classifiers or regressors) models to predict the existence/absence of cropland, such as the nearest neighbour classifier (NN), the naive Bayes classifier (NB), linear regression (LR), ordinary logistic regression (OLR), geographically weighted logistic regression (GWR), classification and regression Trees (CART) (Lesiv et al., 2016). More details on their principles and performance can be found in Lesiv et al. (2016). In contrast, the ranking methods rely on assessing the mapping accuracy and spatial agreement of integrated datasets. The datasets are firstly ranked based on their mapping accuracy. Then, the datasets are integrated, and different types of agreement among datasets are scored. The final map is then created by accumulating the pixels from higher to lower agreement scores until the country’s cropland area or sub-country level reaches close to official agriculture statistics (Fritz et al., 2015; Lu et al., 2020).

Following the integration methods mentioned above, several synergic maps were recently produced at the global scale (Feng & Bai, 2019; Lu et al., 2020; Waldner et al., 2016; Zhang, Ye,
Fang, Li, & Wei, 2019), or mainly for Africa (Fritz et al., 2011; Pérez-Hoyos, Udías, & Rembold, 2020). However, those synergic maps were produced with coarse resolution (≥250 m), which is not adequate for detecting patchy and small agricultural fields in Africa (Whitcraft, Becker-Reshef, & Justice, 2015). Moreover, these maps were initially produced for land cover mapping, where the main target is to accurately detect all land cover types with balanced errors among classes, rather than giving the highest priority to map cropland class accurately. Several researchers already point out that generating a single-type discrete/fraction land cover map could achieve better results than “traditional” multi-class thematic land cover mapping (Nabil et al., 2020; Wu & Zhang, 2017; Xu et al., 2019). Add to that the time inconsistency among input datasets to produce the final synergy map. For example, the synergy map was produced recently by Pérez-Hoyos, Udías, and Rembold (2020) by integrating 20 global, regional, and national remote sensing datasets that were initially produced for inconsistent years (1990-2016). However, the final map was introduced for the nominal year of 2016. This considerable time inconsistency is not acceptable for dynamic land cover classes, such as croplands, varying significantly from a year to another due to the common agriculture shifting practices over Africa (Badiane & Tsitsi, 2014; Ickowitz, 2006).

Recently, four remote sensing landcover datasets, namely ESA-S2-LC20 (CCI, 2017), GFSAD30AFCE (Xiong et al., 2017), FROM-GLC30-2017 (Feng et al., 2018), and CGLS-LC100–2016 (Buchhorn et al., 2020), were produced over Africa in various spatial resolution (20 to 100 meter), covering the period (2015-2017). Despite the datasets’ relatively high spatial resolution, their cropland mapping quality assessment indicated low mapping accuracy and poor spatial consistency (Nabil et al., 2020; Xu et al., 2019). Hence, this study mainly aimed at improving the cropland mapping quality over Africa by integrating these four datasets into a more accurate, up-to-date, and high-spatial-resolution African cropland mask, representing the nominal year 2016. To select the most accurate layer for each agro-ecological zone, the current study also targeted the construction of an extensive set of reference samples through the integration of crowdsource samples, field samples, and photo–interpreted samples for the same period (2015-2017) of the four integrated layers. To assess the quality of final layer, the study also constructed an independent validation set to calculate the final mapping accuracy, in addition to the spatial comparison with two earlier cropland hybrid maps, as well as measuring its consistency with the official statistics.

2. Materials

2.1. Landcover datasets

Four land cover datasets were integrated to build the final cropland mask. The summary of their main characteristics is listed in Table 1. Moreover, the details about each of the four datasets were described in this section.

ESA-S2-LC20 is the sentinel-2 land cover prototype map of Africa produced for 2016 at 20 m by the European Space Agency, mainly to collect users’ feedback for further improvements (http://2016africanlandcover20m.esrin.esa.int/). The cloud-free surface reflectance composites over Africa were derived from more than 30,000 S2A L1C images. Two machine learning classification algorithms were then integrated to generate the final product, with 10 land cover classes and one class for cropland (CCI, 2017). The accuracy was assessed by Lesiv et al. (2017) using two independent datasets, and the overall
Table 1. Landcover datasets that were used for constructing the new cropland layer.

| Dataset            | Sensor          | Resolution (m) | Date          | Classification approach |
|--------------------|-----------------|----------------|---------------|-------------------------|
| ESA-S2-LC20        | Sentinel-2 MSI  | 20             | Dec. 2015 to Dec. 2016 | ML                      |
| GFSAD30AFCE        | Landsat OLI     | 30             | Jul. 2015 to June 2016 | RF & OB                |
| FROM-GLC30-2017    | Landsat OLI     | 30             | Dec. 2015 to Nov. 2016 | RF                     |
| CGLS-LC100-2016    | PROBA-V VNIR    | 100            | Oct. 15 to March 2017 | RF                     |

ML: Machine learning, RF: Random Forest, OB: Object-Based classification.

Accuracy was reported as 65%. Cropland class was massively overestimated at the expense of grasslands, which finally resulted in low cropland user’s accuracy (46% and 50.4%) and producer’s accuracy (71% and 63%).

GFSAD30AFCE is the Global Food Security-support Analysis Data (GFSAD), Cropland Extent Product @ 30-m for Africa developed by Xiong et al. (2017). The product was built by processing 36,924 images (20,214 from Sentinel-2 and 16,710 from Landsat-8) using Google Earth Engine to create 30-m two cloud-free composites (in-season and out-season composites) for entire Africa. In addition to the slope layer, the two seasonal composites were first classified by two Pixel-Based Algorithms (Random Forest & Support Vector Machine) and then integrated into one layer, which was furtherly improved using an Object-Based Algorithm. The final layer has three land cover classes (cropland, water, and non-cropland) and publicly available at the URL: https://lpdaac.usgs.gov/products/gfsad30afcev001/. The reported overall accuracy was 94.5%, while the cropland producer’s accuracy was 85.9%, and the user’s accuracy was reported as 68.5% (Xiong et al., 2017).

FROM-GLC30-2017 is the Finer Resolution Observation and Monitoring of Global Land Cover map for 2017 (FROM-GLC30-2017) produced at high-resolution (30 m) through the integration of multiple datasets (Landsat 8 OLI surface reflectance images, MODIS vegetation index (VI) data (MOD13Q1), a 30-m spatial resolution SRTM DEM, Worldclim, and Night-time light (NTL)). All integrated datasets were furtherly processed to produce seasonal and pheno logical matrices, which were all processed by the random forest classifier to produce the final land cover map with nine classes (available for download at http://data.ess.tsinghua.edu.cn). The same methodology was followed by Feng et al. (2018) to produce FROM-GLC30-2015 over Africa. The reported accuracies for the 2015 map were 75.8% overall accuracy, 68.6% cropland producer accuracy, and 66.6% cropland user accuracy (Feng et al., 2018). However, the accuracy of the 2017 map over Africa has not been reported yet.

CGLS-LC100-2016 is the land cover dataset for 2016 produced by the European Copernicus global land services (CGLS). The dataset has both discrete and fractional land cover layers at 100 m spatial resolution, derived from a 5-days PROBA–V time series imageries, covering the three years (2015-2017). After imaging correction, harmonization, and outliers cleaning, 270 phenological, statistical, temporal, and textural metrics were computed from the PROBA-V 100 m time series. The computed matrices and training samples were inputs for the Random Forest (RF) classifier to produce the final land cover layers (Buchhorn et al., 2020). The resulted land cover discrete layer has 23 land cover classes with only one pure class for croplands. The overall accuracy of the 2016 layer was
reported as 80.4% over Africa (Tsendbazar et al., 2020). The full dataset can be downloaded from the website: https://zenodo.org/record/3518036#.X7u51M0za1s.

2.2. Reference samples

Four crowdsourced and one complementary validation sets were merged into one to cover all African regions with a sufficient number of reference samples (Figure 1). The five merged sets are GFSAD30_VAL, FROM-GLC-VAL, Geo-Wiki cropland database, GVG points, and complementary set. More details on each of the sample sets are provided in this section.

GFSAD30_VAL is the validation dataset that the Global Food Security Analysis-Support Data at 30 Meters (GFSAD30) project, supported by USGS, built to map global cropland at 30 m resolution. The project team collects geographical information on land cover types, crop types, crop intensity, and crop watering type from 2015 onwards. The reference dataset was collected via mobile apps, field visits, and a large part of the data was derived from visual interpretation of high-resolution satellite imageries and time series vegetation indices derived from remote sensing satellite imageries and interpreted by experts. The full dataset can be acquired from the website (https://croplands.org/app/data/search). Another validation dataset belongs to the same (GFSAD30) project built by Xiong et al. (2017) to validate the GFSAD30AFCE product. To build this set, samples were randomly distributed over Africa and then visually interpreted using very high-resolution imageries (sub-meter to 5-meter imageries) provided by the National Geospatial Agency (NGA). The samples located in a 90 square meter frame of homogenous cropland or non-cropland sites were selected as validation samples. The final sample set is available for public use at https://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30AFCE.001. For the current study, both datasets were merged, and the validation points collected during the period (2015-2017) were selected to be time consistent with other land cover datasets. The land cover legend of selected points was

![Figure 1. The five datasets (left) that integrated to build the complete reference samples set for the current study (right).](image-url)
translated into a cropland reference set, with a total number of 2487 samples (938 cropland and 1549 non-cropland samples).

FROM-GLC-VAL is the dataset built by Feng et al. (2018) by dividing the entire globe by the hexagonal scheme, and then 10 samples were assigned randomly in each hexagon to ensure uniformity and objectivity of the distribution of the sample. Land cover types were visually interpreted via seasonal Landsat images, MODIS NDVI time series, and high-resolution images. The dataset is available at http://data.ess.tsinghua.edu.cn/data/temp/. For our current study, the samples only located over Africa were selected and relabeled as cropland/non-cropland (7402 samples: 296 croplands and 7106 non-croplands).

The Geo-Wiki global reference cropland database was produced by Fritz et al. (2017) through a crowdsourcing campaign with the participation of trained students and experts. All participants were asked to revise almost 36,000 possible cropland locations around the world. The participant reviewed all locations via the visual interpreting of high-resolution imageries (Google and Bing) and NDVI time-series profiles from multiple sensors (e.g. Landsat 7 ETM+, Landsat 8 OLI, and MODIS Terra), which cover the 2015–2016 growing season. Out of the entire dataset, the samples located in Africa and confirmed by all participants as 100% cropland were only selected. The derived reference set has 1333 cropland samples.

GVG cropland points were collected using the GVG mobile app during the field trips in 2017 over several African regions. GVG stands for integrated GIS, VIDEO, and GPS system that was mainly built for crop survey. The system can be mounted on a motor and record filed data (as images, locations, and attributes) on moving. Hence, extensive field surveys can be conducted at a convenient time and cost (Tian et al. 2004). All recorded data are memorized into the database server automatically and can be downloaded by users. For the current study, cropland points collected in 2017 were acquired from the data server. The acquired dataset has a total number of 2662 cropland samples.

Complementary points were randomly distributed over the regions with insufficient samples. To check if the point is indeed located in cropland, the NDVI time series profiles of Landsat-8 OLI and Sentinel-2 MSI sensors, in addition to the photointerpretation of recent high-spatial–resolution imageries available on Google Earth Engine (https://earthengine.google.com/), were interpreted. The NDVI time series profiles covered three years (2015-2017), which corresponds to the same epoch of the integrated land cover datasets. Following this approach, 3708 cropland samples were produced.

Finally, to build the complete reference dataset, all the five sets mentioned above were merged after cleaning overlapped points by keeping only one point per location. The final dataset has a total number of 17,592 samples (8937 cropland and 8655 Non-cropland), as previously shown in Figure 1.

### 2.3. Validation dataset

A total of 2561 independent validation samples were systematically generated over Africa (one sample every 100 km). Then, the samples were interpreted using the NDVI time series profiles of Landsat-8 OLI and Sentinel–2 MSI sensors, covering a period of three years (2015-2017), in addition to the photointerpretation of recent high-spatial-resolution imageries available on GEE. Among all African regions with diverse agro-climatic
conditions, two main growing seasons can be identified over the whole continent (July–December for the northern hemisphere and January–June for the southern hemisphere) (Lambert, Waldner, & Defourny, 2016; Waldner et al., 2016; Xiong et al., 2017). Hence, a sample was labeled as cropland if it was cultivated at least one time during the main growing season of any of three years (2015-2017), according to the interpretation of the three years NDVI time series profile, and also was approved as cropland from the interpretation of its pattern using recent high-resolution imageries available on GEE. Following this approach, 507 cropland and 2054 non-cropland samples were produced (Figure 2). Those samples will be used in the independent accuracy assessment of this study’s final constructed layer.

3. Methodology

This work’s methodology is straightforward, and its main steps can be summarized as shown in Figure 3. More details on each step are described in the following section.
3.1. Adopting cropland definition

The cropland definition of all land cover and cropland datasets, as well as the reference data sources addressed in the current study, is similar to FAO’s definition of cropland (Table 2), which includes all agricultural annual standing croplands, cropland fallows, and permanent plantation crops (Feng et al., 2018; Xiong et al., 2017). Only FROM-GLC30-2017 classifies the bare croplands as Bareland, resulting in low mapping accuracy, particularly over regions with a high proportion of fallow fields. In this case, another cropland layer is expected to achieve higher mapping accuracy and thus will be selected to replace FROM-GLC30-2017 in the final constructed layer. More details on selecting the best cropland layer for each zone are provided in the upcoming parts.

Table 2. The cropland definition among multi-data sources addressed in the current study.

|   | FAO               | Agricultural land |   |   |   |   |   |   |   |   |   |
|---|------------------|-------------------|---|---|---|---|---|---|---|---|---|
|   |                  | Cropland          |   |   |   |   |   |   |   |   |   |
|   |                  | Permanent crops   |   |   |   |   |   |   |   |   |   |
|   |                  | Temporary crops   |   |   |   |   |   |   |   |   |   |
|   |                  | Arable land       |   |   |   |   |   |   |   |   |   |
|   |                  | Temporary fallow  |   |   |   |   |   |   |   |   |   |
|   |                  | Pastures          |   |   |   |   |   |   |   |   |   |
|   |                  | Perm. meadows and pastures | | | | | | | | | |
|   |                  | Cultivated        |   |   |   |   |   |   |   |   |   |
|   |                  | Natural growing   |   |   |   |   |   |   |   |   |   |
| CGLS-LC100-2016 | ☑                 | ☑                 | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |
| GFSAD30AFCE      | ☑                 | ☑                 | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |
| FROM-GLC30-2017  | ☑                 | ☑                 | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |
| ESA-S2-LC20      | ☑                 | ☑                 | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |

Figure 3. The flowchart describes the methodology of the current study.
3.2. Deriving cropland layers

All the landcover datasets above-mentioned have one pure class for cropland. Hence, their legend was directly transformed into binary (0 for all non-cropland classes and 1 for cropland). The geographical coordinate system of all datasets was unified to be (“GCS_WGS_1984”, datum: D_WGS_1984). The spatial resolution of cropland layers of both CGLS-LC100-2016 and ESA-S2-LC20 were resampled using the Nearest Neighbour method to 30 m to be at the same resolution of the other two datasets.

3.3. Agro-environmental stratification

In earlier studies (Nabil et al., 2020; Wei et al., 2020), the main factors that limit the accurate mapping of cropland from remote sensing were different among African regions. Thus, one dataset could outperform others in dealing with the region’s limitations and provide a more accurate cropland mapping (Bofana et al., 2020). In this study, the Global Agro-Environmental Stratification (GAES) dataset was used as units of assessing the accuracy of individual cropland layers for each agro-environmental zone (AEZ). The dataset provides four levels of agro-environmental stratification based on the region’s climatic regimes, soil, terrain, elevation conditions, water availability, and land cover proprieties (Mücher et al., 2016). Level 3 was appropriate with the distribution of reference samples set built in the former section. A few adjacent AEZs were merged into one stratum when the number of samples was insufficient. Finally, 41 AEZs were used as units of assessing cropland mapping accuracy (Figure 4).

3.4. Assessing the mapping accuracy

For each AEZ, the error matrix was firstly built, and both the layer’s overall accuracy (OA) and the cropland’s F-score were then estimated. The OA evaluated the overall effectiveness of the classification algorithm (Equation (1)), while the F-score measured the accuracy of cropland class, using the precision (producer accuracy PA) and recall (user accuracy UA) measures (Equations (2)–(4)) (Xiong et al., 2017).

\[
OA = \frac{C_C}{n} \times 100
\]

\[
PA = \frac{S_i}{TS_i} \times 100
\]

\[
UA = \frac{S_i}{TP_i} \times 100
\]

\[
F_{\text{score}} = 2 \times \frac{PA \times UA}{PA + UA}
\]

where \(C_C\) is the total number of correctly classified samples for all classes, \(n\) is the total number of validation samples, \(S_i\) = total number of correctly classified samples for class \(i\); \(TS_i\) = total number of validation samples of class \(i\); \(TP_i\) = total number of classified pixels for class \(i\).
Figure 4. The forty-one agro-environmental zones used by the current study as units of assessing the cropland mapping accuracy by different layers.
Among different estimates of cropland area, by directly counting the number of cropland pixels to the total number of pixels, or after correction based on the information provided by constructed error matrix (e.g. omission and commission errors), the cropland area uncertainty can also be estimated at 95% confidence interval from Equation (5) (Olofsson et al., 2014):

$$Cr_{area} = \bar{x} \pm z \frac{s}{\sqrt{n}}$$

(5)

Where \( \bar{x} \) is the mean cropland area among different estimates, \( z \) is the tabulated Z-value corresponds to the confidence interval level (e.g. \( z = 1.96 \) at 95% confidence interval, \( s \) is the standard deviation, and \( n \) is the number of estimates.

### 3.5. Constructing the final mask

For each AEZ, the cropland layer with the highest overall accuracy and cropland F-score was identified as the best layer. The final cropland mask was constructed by merging all identified layers over Africa.

### 3.6. Applying urban masking

The constructed cropland layer was furtherly improved by masking out urban areas that were wrongly classified as cropland by individual cropland layers. The urban mask over Africa was derived from the Global Urban Footprint (GUF) layer. The layer is a binary urban mask constructed from the processing of thousands of TerraSAR-X/TanDEM-X images covering the period (2011–2013) at the global scale. More details of the GFU prepossessing and specifications can be found in Esch et al. (2013). The urban mask is available at 12 and 84 m spatial resolutions. For the current study, the finer urban mask (at 12 m) was used after upscaling to 30 m using the Nearest Neighbour method to be spatially consistent with the cropland mask.

### 3.7. Assessing the quality of the final map

The newly constructed mask’s quality was also assessed through the spatial comparison with two earlier cropland synergy maps. The first is the Unified Cropland Layer (UCL) that was produced by Waldner et al. (2016) for the nominal year 2014 through the integration of 49 global, regional, and local Land cover datasets covering the period (1990-2014). The integration between datasets was based on multi-criteria analysis. Each dataset got a score according to its adequacy with the legend and the spatial resolution, timeliness, and confidence level. The dataset that got the highest score over a region was then selected to be a part of the unified cropland layer (more details can be found in Waldner et al. (2016)). The final global cropland layer has a spatial resolution of 250 m and an overall accuracy of 82% to 95%, as reported by the layer producers. The second synergy map is the JRC layer developed by Pérez-Hoyos et al. (2020) for Africa by integrating 21 global, African, and local Land cover datasets. The integration was done by selecting the best dataset after evaluation at the country level, according to five criteria: timeliness, spatial resolution, comparison with FAO statistics, accuracy assessment, and expert
evaluation. The final layer was in 250 m spatial resolution for the nominal year 2016 and can be acquired upon request from the dataset’s producers.

Besides the spatial comparison among synergic cropland layers, their closeness to FAO cropland official statistics at the country level was also measured. The cropland statistics were acquired from the website (http://www.fao.org/faostat/en/#data/RL) of the Food and Agriculture Organization of the United Nations (FAO) Land Use database (FAOSTAT). These statistics were compiled from various sources (e.g. national statistics, agricultural census, household surveys, and expert opinions) and were widely used in various agricultural management and production forecasts and also as an essential component of synergic mapping studies (Chen et al., 2019; Clinton et al., 2015; Lu et al., 2017; Tsendbazar et al., 2015). Moreover, FAO land-use dataset separates cropland from pastures and provides detailed statistics for the three cropland subcategories (temporary crops, temporary fallow, and permanent crops). Following our adopted cropland definition, the cropland areas, including the three cropland’s subcategories, for all African countries were derived from the full land-use dataset. Since the satellite-based products addressed in the current study cover the period from 2015 to 2017, the average cropland area for the same period of three years was calculated to provide more accurate and stable estimates for the nominal year of 2016.

4. Results

4.1. The best cropland layer per AEZ

Based on the validation of all cropland layers against the reference sample set (Figure S1), the most accurate cropland layer, with the highest cropland F-score and overall accuracy, was identified for each AEZ (Figure 5). The GFSAD30AFCE was the most accurate cropland layer over the majority of African zones (25 zones include the Southern, Eastern, Central Africa, and western coast), followed by CGLS-LC100-2016 (9 zones cover North-Western Africa), FROM-GLC30-2017 (6 zones cover Morocco, north Algeria, and Egypt), and ESA-S2-LC20 (only one AEZ over Ethiopia). Moreover, the AEZs level analysis indicated that the cropland F-score of at least one out of the four layers was falling below 0.50 over almost half (20) of AEZs, indicating poor cropland mapping accuracy (Figure S1). This highlights the importance of the current methodology of validating multiple cropland layers to select the best layer for each zone to build the final layer with improved mapping accuracy.

4.2. The final cropland mask layer

By spatially joining the best cropland layers, identified in Figure 5, the final mask was constructed as a binary layer with a 30 m spatial resolution representing the African cropland distribution for the nominal year 2016 (Figure 6). The newly constructed layer mapped cropland area as 282 Mha, representing 9.36% of the African continent area. Moreover, the United Nations classifies Africa into five main sub-regions. By overlaying with the cropland layer, Eastern Africa has 36.2% of African cropland, followed by Western Africa (28.6%). The Southern Africa region has the lowest cropland area (8.5%) among the other five main regions (Figure 6).
4.3. Data records

The layer constructed by the current study represents the geographical distribution of Africa cropland in 30 m spatial resolution. The cropland class is representing arable lands, permanent crops, and fallow lands. Since the input remote sensing land cover datasets covered the period (2015-2017), the constructed layer was introduced for the nominal year of 2016. The final layer’s legend has three categories (0: Background, 1: Non-cropland, and 2: Cropland). Besides the primary layer (in tiff format), there are sex supplementary layers (in TFW, XML, OVR, CPG, and DBF formats) that help users to easily open and handle the cropland layer through the common geographic information system (GIS) software, like ArcMap and QGIS. All these layers were stored in one dataset for open sharing publicly; see the “Data Availability Statement” section. Additional data characteristics that could be relevant for users are provided in Table 3.

Figure 5. The most accurate cropland layer was identified for each Africa agro-ecological zone (AEZ).
Figure 6. The African cropland mask built by the current study at 30 m spatial resolution for 2016, and the cropland percentage of the five main African sub-region.

Table 3. Data characteristics of the African cropland mask 2016.

| Data Characteristic    | Description                      |
|------------------------|----------------------------------|
| Data format            | GeoTiff                          |
| Epoch                  | 2016                             |
| Coordinate system      | GCS_WGS_1984                     |
| Size                   | 1.68 GB                          |
| Data type              | Unsigned integer 8bit            |
| Compression type       | LZW compression                  |
| No Data value          | 0                                |
| Value                  | 1: Non-cropland, 2: Cropland     |

Table 4. The accuracy assessment report of the constructed cropland layer.

| Reference | NC    | C     | Total | OA%   | Cropland PA% | Cropland UA% | Cropland F-Score | Non-cropland PA% | Non-cropland UA% | Non-cropland F-Score |
|-----------|-------|-------|-------|-------|--------------|--------------|------------------|------------------|------------------|---------------------|
| NC        | 2033  | 193   | 2226  | 91.64 | 61.93        | 93.73        | 0.75             | 98.98            | 91.33            | 0.95                |
| C         | 21    | 314   | 335   |       |              |              |                  |                  |                  |                     |
| Total     | 2054  | 507   | 2561  |       |              |              |                  |                  |                  |                     |

C: cropland, NC: Non-cropland, OA: overall accuracy, PA: producer accuracy, and UA: user accuracy.
5. Technical validation

5.1. Accuracy assessment

The final cropland layer achieved an overall accuracy of 91.6% against the validation samples set built for the current study (Table 4). Although the high achieved cropland user accuracy (93.7%) and F-score (0.75), the producer accuracy was lower (61.9%). This indicated that the newly constructed layer tends to wrongly omit several cropland areas (high omission error corresponds to low producer accuracy) than to commit.

The mapped area of cropland was 9.38%, as directly estimated from the cropland pixels’ proportion to total pixels. However, this direct estimation is subject to classification errors (Wu & Li, 2004; Devaser & Luhach, 2016). Instead, the error matrix (Table 4) indicated that the omission error was 38.07% and the commission error was 6.27%, with a net error of 31.8% omission (38.07 - 6.27). By correcting the mapped area based on the error matrix, as proposed by (Czaplewski & Catts, 1992; Dong, Liu, Wang, Chen, & Gallego, 2017), the corrected area is 12.36% (9.38 + (9.38 × 31.8/100)). Another cropland area could be estimated as 19.80% from the number of reference cropland samples to total cropland samples (507/2561). However, the last two estimates assuming the golden samples representability where samples ratio represent the ratio between cropped and non-cropped areas, and the mislabeling errors in interpreted samples are absent. The two assumptions are hard to ensure at such a continental scale. Among the three estimates of cropland area, the uncertainty of area estimation can be estimated (from Equation (5)) as 13.8 (95% CI 7.72 to 19.9).

5.2. Spatial comparison

The spatial comparison between unified cropland layer UCL, produced by Waldner et al. (2016), for 2014 and newly constructed cropland mask (CLM) for 2016 indicated a significant difference in cropland distribution, particularly over the Sahel, Central, and western African regions (Figure 7(a)). The UCL appeared to be overestimating cropland distribution over vast African regions (9.17% of Africa area, green colour in Figure 7(a)) compared to CLM. The visual interpretation of recent and high-resolution satellite images on Google Earth Engine (GEE) showed that the two layers’ discrepancies were mostly clustered over regions mainly dominated by forest and shrubs, with small and scattered agricultural fields. Over these regions, the whole area (includes agricultural fields, forest, and shrubs) was wrongly classified as cropland by UCL, while the newly constructed CLM did not detect the patchy fields (e.g. in northwestern Zambia Figure 7(c)). This could be due to the difficulty of mapping cropland in high fragmented and patchy agricultural systems by the four integrated cropland layers. Moreover, the cropland area mapped by UCL was 1.34 times the area mapped by CLM over Africa (UCL area = 377.84 Mha, CLM area = 281.89 Mha as reported in Table S1), while it was estimated as only 277.38 Mha by FAOSTAT.

On the contrary, large cropland areas were missing from UCL but successfully detected by CLM, such as large orchards regions in Kairouan, Tunisia (Figure 7(b)). This could be attributed to the inconsistency of cropland definition among datasets used to produce UCL. In other words, the single cropland layer selected over the region to build the final UCL might not consider the orchards as a part of the cropland area. While in our CLM, croplands include permanent crops such as orchards.

The comparison between CLM and JRC layer, produced by Pérez-Hoyos et al. (2020), revealed relatively higher agreement than with UCL. The area of disagreement with JRC was
only 6% (areas in green and magenta Figure 8(a)), compared to 12.3% with UCL. Compared to our CLM, the JRC layer overestimated cropland, particularly in the Sahel region, where vast regions of bare and sparse vegetation were wrongly classified as cropland (Figure 8(b)). Moreover, large agricultural fields near the Congo River were missing from the JRC layer, while high-resolution satellite images on GEE show large slashed and burned areas, a common practice in Africa to clear the fields before re-cultivation (Figure 8(c)).

Overall, the spatial comparison of the newly constructed mask with the two earlier synergy maps revealed the improvement in the quality of mapping African cropland at 30 m spatial resolution over vast regions but also indicated that a large number of patchy agricultural fields, mainly in central Africa, were missing from the newly constructed mask. However, this could be improved by the integration of upcoming cropland datasets.

5.3. The consistency with FAO official statistics

The African countries’ cropland area estimated from the newly constructed map and the two earlier cropland synergy maps were compared to FAOSTAT areas (Table S1). According to FAO, Africa’s cropland area is about 277 Mha, as average estimates between 2015 and 2017, representing 9.23% of Africa’s total area. Both UCL and JRC seem to be overestimating the Africa cropland area when compared with the numbers reported by FAO. In contrast, our new CLM estimated the cropland area as 282 Mha, closer to the FAOSTAT area than the other two maps (UCL = 377.84 Mha and JRC = 320.02 Mha).
Figure 8. (a) The spatial comparison between the JRC cropland layer and the cropland mask (CLM) built by the current study over Africa and highlighting two locations (b & c) as an example of the spatial discrepancies among two cropland layers.
Figure 9. The scatter plots show the relationship between the cropland areas of 49 African countries (in Mha) estimated from the three cropland layers compared to FAOSTAT areas.
Despite the significant difference in cropland areas among layers, the three cropland synergic layers showed high agreement with FAO official statistics at the countries level (high $R^2 > 0.85$ for the three layers, Figure 9). However, cropland areas estimated from the CLM have the lowest root mean square error (RMSE = 2.85 Mha) compared to FAOSTAT. This indicates that the newly constructed cropland mask is relatively more consistent with official statistics than the other two layers.

6. Data set value

With the availability of several cropland layers for Africa derived from remote sensing data, this research produced a more accurate cropland layer for the Africa continent at 30 m spatial resolution. Unlike the earlier synergy maps that were initially produced from integrating many time-inconsistent layers, the new layer is expected to better represent Africa’s actual cropland distribution for the years 2015-2017. The users could refer to the current study to select the most accurate cropland layer for the region of their interest. Considering its higher spatial resolution than earlier studies’ synergy maps (usually 250 m), the 30 m constructed layer could be used as an initial mask layer for agriculture and environmental monitoring projects. Moreover, the layer can help in producing cropland samples due to its low commission error.

Acknowledgments

The first author, Mohsen Nabil, acknowledges the Chinese Academy of Sciences (CAS) for supporting him in carrying out the research. The authors are very grateful to all colleagues in the CropWatch team at AIRCAS for providing the GVG samples used in the current study. We are also very grateful to the journal’s editor, co-editor and anonymous reviewers for their detailed reviews that helped us finalizing the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences [XDA19030201], and National Natural Science Foundation of China [41861144019 and 41561144013].

Notes on contributors

Mohsen Nabil is an agricultural remote sensing specialist at the national authority for remote sensing and space sciences (NARSS), Egypt. He is currently pursuing his PhD degree at the State Key Laboratory of Remote Sensing Science, AIRCAS. He received his BSc degree in agriculture engineering from Cairo University, Egypt, in 2011 and his MSc degree in Geo-informatics and environmental management in 2016 from the Mediterranean Agronomic Institute of Chania (MAICh), Greece. He is currently a team member of CropWatch. His research is mainly focused on remote sensing of agriculture, especially in cropland and land cover mapping, change trend analysis, crop type classification, and area estimation. He has authored and co-authored several peer-reviewed publications devoted to remote sensing applications in agriculture in Africa.
**Dr. Miao Zhang** is an associate professor at the State Key Laboratory of Remote Sensing Science, AIRCAS. He received his Ph.D. in Cartography and Geographical Information System from UCAS in 2014. Since then, he has worked as a key member of the CropWatch team. He was nominated as an Executive Committee member of GEOGLAM Flagship and one of the co-leads of the Joint Experiments for Crop Assessment and Monitoring (JECAM) Initiative, which strengthened the internationally collaborative relationships with researchers from more than 20 countries and other international organizations under the GEOGLAM context. His research focuses on remote sensing-based agriculture monitoring. More specifically, his work mainly includes (1) crop classification using multi-source remote sensing imageries; (2) crop area estimation and early warning; and (3) quantitative approaches for crop condition assessment and yield prediction by the integration of remote sensing and weather data. He got more than 60 scientific papers since 2010, and his research outputs span across various disciplines, including remote sensing methodology, water resource, and agronomy.

**Prof. Dr. Bingfang Wu** is the deputy director of the Center for Chinese Agricultural Policy, CAS, and is a Professor in the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences (AIRCAS). He is co-chair of GEO/GEOS GEGLAM flagship since 2009; member of CAS education committee since 2013; Chair of Digital Agriculture Technical Committee, China National Commission of International Society of Digital Earth; Chair of Remote Sensing for Ecology Technical Committee, China Ecology Association. Prof. Wu is widely recognized as one of the world’s leading experts in crop Monitoring, food security analysis, and water resources management. His research interests are remote sensing methodologies and applications in agriculture, water resource, and ecosystem, especially in monitoring crops, evapotranspiration (ET), and drought. He and his team have developed remote sensing-based operational monitoring systems with novel technologies as CropWatch since 1992, DroughtWatch since 1999, ETWatch since 2002, which have been customized for and deployed at many stakeholders. Prof. Wu has published three monographs and more than 300 peer-reviewed papers. He is the Editor-in-Chief of China LandCover Atlas (1:1,000,000).

**Jose Bofana** is currently a Doctoral Student at AIRCAS. Mr. José has substantial experience in GIS and Remote Sensing as well as Agriculture and Environment Engineering. Mr. José’s areas of interest are Land use and Land cover assessment, Environmental Science, Agronomy, and Agricultural Plant Science. He worked as coordinator of the Pilot Project in Agriculture Information Management and Precision Agriculture in the AgIM program carried out by the Catholic University of Mozambique. He also worked as coordinator of the Centre for Agricultural and Sustainable Development Research (CIADS) at the Faculty of Agricultural Sciences, Catholic University of Mozambique. Currently, his research areas are engaged in the following specialties: geospatial analysis, land cover, and land-use change assessment, precision farming, satellites, and drone data collection and analysis.

**Abdelrazek Elnashar** received his BSc degree in Agriculture Sciences from Kafrelsheikh University, Egypt, and MSc in Agriculture sciences from Cairo University, Egypt. Currently, he is pursuing his PhD degree at the State Key Laboratory of Remote Sensing Science, AIRCAS. As a CropWatch team member at AIRCAS, He has worked on cropping intensity and crop mapping and modeling using remote sensing, machine learning, and cloud computing. His research also includes cropland dynamics and its association with water availability in the East Africa region.
ORCID

Mohsen Nabil  
http://orcid.org/0000-0003-2362-6711

Jose Bofana  
http://orcid.org/0000-0002-2534-450X

Abdelrazek Elnashar  
http://orcid.org/0000-0001-8008-5670

Data availability statement

The data that support the findings of this study are openly available in the Science Data Bank (ScienceDB) at http://www.dx.doi.org/10.11922/sciencedb.j00076.00063, and figshare at https://doi.org/10.6084/m9.figshare.13520141.v1.

Open scholarship

This article has earned the Center for Open Science badge for Open Data. The data are openly accessible at http://www.dx.doi.org/10.11922/sciencedb.j00076.00063, and figshare at https://doi.org/10.6084/m9.figshare.13520141.v1.

References

Badiane, O., & Tsitsi, M. (2014). The theory and practice of agriculture, growth, and development in Africa (WIDER Working Paper, 2014/061). doi:10.35188/unu-wider/2014/782-0

Bofana, J., Zhang, M., Nabil, M., Wu, B., Tian, F., Liu, W., … Moyo, C. (2020). Comparison of different cropland classification methods under diversified agroecological conditions in the Zambezi River Basin. Remote Sensing, 12(13), 2096.

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., & Smets, B. (2020). Copernicus global land cover layers—collection 2. Remote Sensing, 12(6), 1044.

CCI, L. C. (2017). S2 prototype land cover 20 m map of Africa.

Chen, D., Lu, M., Zhou, Q., Xiao, J., Ru, Y., Wei, Y., & Wu, W. (2019). Comparison of two synergy approaches for hybrid cropland mapping. Remote Sensing, 11(3), 213.

Clinton, N., Le, Y., & Peng, G. (2015). Geographic stacking: Decision fusion to increase global land cover map accuracy. ISPRS Journal of Photogrammetry and Remote Sensing, 103, 57–65.

Czaplewski, R. L., & Catts, G. P. (1992). Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sensing of Environment, 39(1), 29–43.

Devaser, V., & Luhach, A. K. (2016). An approach to analyse the agriculture acreage and estimate production. Indian Journal of Science and Technology, 9(28), 28.

Dong, Q., Liu, J., Wang, L., Chen, Z., & Gallego, J. (2017). Estimating crop area at county level on the North China plain with an indirect sampling of segments and an adapted regression estimator. Sensors, 17(11), 2638.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., … Dech, S. (2013). Urban footprint processor—fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. IEEE Geoscience and Remote Sensing Letters, 10(6), 1617–1621.

Feng, D., Yu, L., Zhao, Y., Cheng, Y., Xu, Y., Li, C., & Gong, P. (2018). A multiple dataset approach for 30-m resolution land cover mapping: A case study of continental Africa. International Journal of Remote Sensing, 39(12), 3926–3938.

Feng, M., & Bai, Y. (2019). A global land cover map produced through integrating multi-source datasets. Big Earth Data, 3(3), 191–219.
Fritz, S., Linda, S., Liangzhi, Y., Chris, J., Inbal, B. R., Lieven, B., & Renato, C. (2013). The need for improved maps of global cropland. *Eos, Transactions American Geophysical Union*, 94(3), 31–32.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., … Obersteiner, M. (2015). Mapping global cropland and field size. *Global Change Biology*, 21(5), 1980–1992.

Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., … Laso-Bayas, J.-C. (2017). A global dataset of crowdsourced land cover and land use reference data. *Scientific Data*, 4(1), 170075.

Fritz, S., You, L., Bun, A., See, L., McCallum, I., Schill, C., … Obersteiner, M. (2011). Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophysical Research Letters*, 38(4). doi:10.1029/2010gl046213

Frolking, S., Qiu, J., Boles, S., Xiao, X., Liu, J., Zhuang, Y., … Qin, X. (2002). Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China. *Global Biogeochemical Cycles*, 16(4), 38–31–38-10. doi:10.1029/2001gb001425

FSIN. (2020). *Food security information network, 2020 global report on food crises: Joint analysis for better decisions*. Rome, Italy and Washington, DC: Food and Agriculture Organization (FAO); World Food Programme (WFP); and International Food Policy Research Institute (IFPRI). Retrieved from https://www.fsinplatform.org/global-report-food-crisis-2020.

Hao, P.-Y., Tang, H.-J., Chen, Z.-X., Yu, L., & Wu, M.-Q. (2019). High resolution crop intensity mapping using harmonized Landsat-8 and Sentinel-2 data. *Journal of Integrative Agriculture*, 18(12), 2883–2897.

Ickowitz, A. (2006). Shifting cultivation and deforestation in tropical Africa: Critical reflections. *Development and Change*, 37(3), 599–626.

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(3), 232.

Lesiv, M., Fritz, S., McCallum, I., Tsendbazar, N., Herold, M., Pekel, J.-F., … Van De Kerchove, R. (2017). Evaluation of ESA CCI prototype land cover map at 20m. IIASA Working Papaer, WP-14-021.

Lesiv, M., Moltchanova, E., Schepaschenko, D., See, L., Shvidenko, A., Comber, A., & Fritz, S. (2016). Comparison of data fusion methods using crowdsourced data in creating a hybrid forest cover map. *Remote Sensing*, 8(3), 261.

Lu, M., Wenbin, W., You, L., See, L., Fritz, S., Yu, Q., … Bing, X. (2020). A cultivated planet in 2010: 1. The global synergy cropland map. *Earth System Science Data*, 12(3), 1913–1928.

Lu, M., Wu, W., You, L., Chen, D., Zhang, L., Yang, P., & Tang, H. (2017). A synergy cropland of China by fusing multiple existing maps and statistics. *Sensors*, 17(7), 7.

Mücher, S., De Simone, L., Kramer, H., De Wit, A., Roupioz, L., & Hazeu, G. (2016). A new Global Agro-Environmental Stratification (GAES) (2761). Wageningen Environmental Research. http://library.wur.nl/WebQuery/wurpubs/fulltext/400815 (accessed on 20 October 2020).

Mutanga, O., Dube, T., & Galal, O. (2017). Remote sensing of crop health for food security in Africa: Potentials and constraints. *Remote Sensing Applications: Society and Environment*, 8, 231–239.

Nabil, M., Zhang, M., Bofana, J., Wu, B., Stein, A., Dong, T., … Shang, J. (2020). Assessing factors impacting the spatial discrepancy of remote sensing based cropland products: A case study in Africa. *International Journal of Applied Earth Observation and Geoinformation*, 85, 102010.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42–57.

Ozdogan, M., Yang, Y., Allez, G., & Cervantes, C. (2010). Remote sensing of irrigated agriculture: Opportunities and challenges. *Remote Sensing*, 2(9), 2274–2304.

Pérez-Hoyos, A., Rembold, F., Kerdelles, H., & Gallego, J. (2017). Comparison of global land cover datasets for cropland monitoring. *Remote Sensing*, 9(11), 1118.

Pérez-Hoyos, A., Udías, A., & Rembold, F. (2020). Integrating multiple land cover maps through a multi-criteria analysis to improve agricultural monitoring in Africa. *International Journal of Applied Earth Observation and Geoinformation*, 88, 102064.
Tian, Y., Bingfang, W., Wenbo, X., Jianxi, H., & Wenting, X. (2004). An effective field method of crop proportion survey in China based on GVG integrated system. IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium, 6, 4028–4030. Anchorage, AK, USA.

Tsendbazar, E., Tarko, A., Linlin, L., Herold, M., Lesiv, M., Fritz, S., & Victor, M. (2020, September). Copernicus global land service: Land cover 100m: Version 3 globe 2015-2019: Validation report. Geneve, Switzerland: Zenodo.

Tsendbazar, N. E., De Bruin, S., Fritz, S., & Herold, M. (2015). Spatial accuracy assessment and integration of global land cover datasets. Remote Sensing, 7(12), 15804–15821.

Wu, B., and Li, Q (2004). Crop area estimation using remote sensing on two-stage stratified sampling. Crop Area Estimation Using Remote Sensing on Two Stage Stratified Sampling, In Proceedings of the XXth International Society for Photogrammetry and Remote Sensing (ISPRS) congress: Geo-information Bridging Continents, July 12–23, 2004, Istanbul, Turkey.

Waldner, F., De Abelleyra, D., Verón, S. R., Zhang, M., Wu, B., Plotnikov, D., … Kussul, N. (2016). Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity. International Journal of Remote Sensing, 37(14), 3196–3231.

Waldner, F., Fritz, S., Di Gregorio, A., Plotnikov, D., Bartalev, S., Kussul, N., … Defourny, P. (2016). A unified cropland layer at 250m for global agriculture monitoring. Data, 1(1), 3.

Wei, Y., Lu, M., Wu, W., & Ru, Y. (2020). Multiple factors influence the consistency of cropland datasets in Africa. International Journal of Applied Earth Observation and Geoinformation, 89, 102087.

Whitcraft, A., Becker-Reshef, I., & Justice, C. (2015). A framework for defining spatially explicit earth observation requirements for a global agricultural monitoring initiative (GEOGLAM). Remote Sensing, 7(2), 1461–1481.

Wu, B., & Zhang, M. (2017). Remote sensing: Observations to data products. Acta Geographica Sinica, 72(11), 2093–2111.

Xiong, J., Thenkabail, P. S., Tilton, J. C., Gumma, M. K., Teluguntla, P., Yadav, K., … Gorelick, N. (2017). 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. Remote Sensing, 9(10), 1065.

Xu, Y., Yu, L., Peng, D., Li, C., Huang, X., … Gong, P. (2019). Comparisons of three recent moderate resolution African land cover datasets: CGLS-LC100, ESA-S2-LC20, and FROM-GLC-Africa30. International Journal of Remote Sensing, 40(16), 6185–6202.

Xu, Y., Yu, L., Zhao, F. R., Cai, X., Zhao, J., Lu, H., & Gong, P. (2018). Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa. Remote Sensing of Environment, 218, 13–31.

Yu, Q., Liangzhi, Y., Ulrike, W.-s., Yating, R., Alison, K. J., Steffen, F., … Peng, Y. (2020). A cultivated planet in 2010: 2. The global gridded agricultural production maps. Earth System Science Data, 12 (4), 3545–3572.

Zhang, C., Ye, Y., Fang, X., Li, H., & Wei, X. (2019). Synergistic modern global 1 Km cropland dataset derived from multi-sets of land cover products. Remote Sensing, 11(19), 2250.