

Article

A High-Resolution Cropland Map for the West African Sahel Based on High-Density Training Data, Google Earth Engine, and Locally Optimized Machine Learning

Kaboro Samasse ^{1,2,*}, Niall P. Hanan ³ , Julius Y. Anchang ³  and Yacouba Diallo ²¹ Geospatial Sciences Center of Excellence, South Dakota State University, Brookings, SD 57007, USA² IPR/IFRA, BP 06 Koulikoro, Mali; yacouba.diallo@ipr-ifra.edu.ml³ Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM 88003, USA; nhanan@nmsu.edu (N.P.H.); anchang@nmsu.edu (J.Y.A.)

* Correspondence: kaboro.samasse@jacks.sdstate.edu

Received: 23 March 2020; Accepted: 28 April 2020; Published: 1 May 2020



Abstract: The West African Sahel Cropland map (WASC30) is a new 30-m cropland extent product for the nominal year of 2015. We used the computing resources provided by Google Earth Engine (GEE) to fit and apply Random Forest models for cropland detection in each of 189 grid cells (composed of 100 km², hence a total of $\sim 1.9 \times 10^6$ km²) across five countries of the West African Sahel (Burkina Faso, Mauritania, Mali, Niger, and Senegal). Landsat-8 surface reflectance (Bands 2–7) and vegetation indices (NDVI, EVI, SAVI, and MSAVI), organized to include dry-season and growing-season band reflectances and vegetation indices for the years 2013–2015, were used as predictors. Training data were derived from an independent, high-resolution, visually interpreted sample dataset that classifies sample points across West Africa using a 2-km grid ($\sim 380,000$ points were used in this study, with 50% used for model training and 50% used for model validation). Analysis of the new cropland dataset indicates a summed cropland area of $\sim 316 \times 10^3$ km² across the 5 countries, primarily in rainfed cropland (309×10^3 km²), with irrigated cropland area (7×10^3 km²) representing 2% of the total cropland area. At regional scale, the cropland dataset has an overall accuracy of 90.1% and a cropland class (rainfed and irrigated) user's accuracy of 79%. At bioclimatic zones scale, results show that land proportion occupied by rainfed agriculture increases with annual precipitation up to 1000 mm. The Sudanian zone (600–1200 mm) has the highest proportion of land in agriculture (24%), followed by the Sahelian (200–600 mm) and the Guinean (1200 +) zones for 15% and 4%, respectively. The new West African Sahel dataset is made freely available for applications requiring improved cropland area information for agricultural monitoring and food security applications.

Keywords: agricultural land area; Sahel; West Africa; machine learning; Earth Engine

1. Introduction

Timely and accurate information on cultivated areas is of paramount importance for food security planning [1,2]. This is particularly true in developing regions, like the West African Sahel, where most cropland is rainfed and agricultural production is susceptible to fluctuations in precipitation [3]. Earth Observation (EO) satellites can contribute significantly to providing information to the agricultural sector, as they allow for consistent land surface imaging over broad spatial extents (regionally or globally) with high revisit frequency [4]. That makes these technologies suitable for monitoring vegetation [5], cropland area [6–8], and agricultural production [9–12]. Optical remote sensing in

particular offers unique possibilities for mapping cropland extent, in addition to monitoring the growth and eventual yield of cultivated lands [13,14].

The accuracy of remote-sensing-based land cover (including cropland) products varies considerably depending on the scale of assessment, the statistical approaches adopted, and the quality and quantity of training and evaluation data. Samasse et al. [15] recently reviewed eight global and regional land cover maps [8,16–22] using high-density evaluation data for the five countries of the Western Sahel (Burkina Faso, Mali, Mauritania, Niger, and Senegal). The study focused uniquely on cropland classes. They found large errors in all existing products, particularly in the coarser resolution (>300 m) products. However, even the higher resolution (~30 m) datasets had accuracy statistics (“user’s accuracy”) of less than 75%, and all existing products were greatly biased towards overestimating the area of active cropland in the region. More recent studies benefitted from high spatial resolution data (10 m or less) to map cropland in the region. For example, Tong et al. [23] used full-year Sentinel-2 NDVI data and Random Forest classifiers to separate cropland from fallow across the Sahel belt [23] at 10 m resolution, reporting an overall average accuracy of 88% for crop and fallow classes. However, they also used several land cover products with known moderate or low accuracy for cropland extent to develop the fallow/cropland map. Specifically, the CGLS LCC 100 m [24], and ESA CCI 300 m [18] maps, used as croplands mask in Tong et al. [23], have low cropland class-specific accuracy (~60%) [25] and high area overestimation [15], respectively. That may lead, via error propagation, to important misclassifications in the final product, attenuating our ability to retrieve cultivated land area as a precursor to yield modeling and prediction.

The clear need for improvements in cropland area assessments in the Sahel region, coupled with the potential for improvements made possible by using higher resolution data, also increases the need for computational resources, new methods, and technical skills for effective processing and analysis. Google Earth Engine (GEE) [26] is one of the platforms currently facilitating access and processing of larger data volumes for diverse operational applications including cropland mapping. The Landsat data archive in particular, with 30 m spatial resolution, long temporal record, and no cost, provides an opportunity to map large scale agricultural regions consistently and in greater detail [27,28]. Recent satellite instrument additions (e.g., Copernicus Sentinel Instruments) provide increasing opportunities to combine data from multiple sources for improved spatial, temporal, and radiometric resolution.

In this study, we leverage the availability of more than 400,000 land-cover training data points for the year 2013 [15,29], with hundreds of cloud-free Landsat-8 images (for the years 2013–2015), to train locally-optimized Random Forest models predicting presence and absence of rainfed and irrigated agricultural fields across the non-desert (MAP > 200 mm/y) land area of the West African countries of Mauritania, Senegal, Mali, Burkina Faso, and Niger. Our analysis grid is composed of 267 (100 × 100 km) grid squares, each processed separately using Google Earth Engine (GEE) to fit and apply locally optimized Random Forest models for cropland detection at 30 m. We analyze our results to estimate accuracy and uncertainty of the new classification, present summary statistics for cropland in the region, and make the new West African Sahel Cropland dataset (under the name WASC30) freely available for applications requiring improved crop area data for agricultural monitoring and food security.

2. Materials and Methods

2.1. Reference Data

Reference data on presence and absence of rainfed and irrigated agriculture were obtained from the Rapid Land Cover Mapper (RLCM) [29–31] for the year 2013. The RLCM approach uses local experts and visual interpretation of 30 m Landsat images to assess land cover type, sampled at 2 km intervals across West Africa [29,30,32]. While RLCM data are available for several epochs (1975, 2000, 2013), we use only the 2013 data as training data for this study. The dataset provides classification into one of 25 land cover types for each centroid of the 2 km grid, with possible land cover classes including

multiple non-agricultural classes, and agricultural classes including rainfed and irrigated cropland. The approach, based on expert visual interpretation, with specific local knowledge of the environments being classified, is expected to show better results than semi- or fully automated classifiers, particularly for the cropland class across West Africa [32].

Quality control for the reference data was carried out using multiple sources of ancillary data, including thousands of aerial photographs taken by the USGS team, high-resolution verification using Google Earth satellite imagery, and field validation in each country, facilitating systematic verification of land cover assessments [30]. In addition, image interpretation and land cover assessments carried out by national experts were reviewed and revised during regular collaborative workshops in West Africa, to ensure consistent practice between country teams and USGS partners. Further details are provided by Samasse et al. [15].

In this study, we regrouped the 25 land cover classes into 3 classes (rainfed and irrigated agriculture and non-agricultural) and used 50% of the 2 km by 2 km data points for year 2013 as reference information for training the classification algorithm and the other 50% for assessing the classified product. Reduced data-density in some areas (e.g., on the coastal and desert margins) resulted in a total of 383,464 reference data points (non-crop, rain-fed, and irrigated classes) across our West Africa study domain.

2.2. Google Earth Engine (GEE)

Google Earth Engine is a cloud-based platform for regional and planetary scale earth observation data retrieval and processing. Its advantage is to store the petabytes of freely available data (e.g., Landsat imagery) in the cloud, avoiding the need for data download, while providing high-performance parallel computing resources to process large datasets [26]. GEE thus facilitates computationally cumbersome geospatial analysis with minimal local computing and storage resources. GEE makes use of an application programming interface (API in JavaScript or Python), allowing for data processing and visualization at different scales. The GEE platform also implements several Machine Learning algorithms (Support Vector Machine, Random Forest) known to be effective for land cover and land use classification in general and cropland mapping in particular [8,33–35].

2.2.1. Landsat 8 Surface Reflectance (SR)

The Landsat mission is a joint initiative of the USGS and NASA providing consistent earth observation data at sub-100 m spatial resolution since the 1970s. Surface reflectance data from the Landsat-8 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) is available in GEE from April 2013 to present. Table 1 contains information on the Landsat-8 SR Tier 1 collection spectral bands used in this study.

Table 1. Landsat 8 band description and wavelengths. The “pixel_qa” band provides metadata on scene quality such as cloud cover for each pixel.

Name	Band description	Wavelength (μm)
B2	Band 2 (blue) surface reflectance	0.452–0.512
B3	Band 3 (green) surface reflectance	0.533–0.590
B4	Band 4 (red) surface reflectance	0.636–0.673
B5	Band 5 (near infrared) surface reflectance	0.851–0.879
B6	Band 6 (shortwave infrared 1) surface reflectance	1.566–1.651
B7	Band 7 (shortwave infrared 2) surface reflectance	2.107–2.294
pixel_qa	Pixel quality attributes generated from the CFMASK algorithm.	—

(Source: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_SR).

2.2.2. Vegetation Indices

In addition to the individual band reflectances, remote sensing derived vegetation indices (VI) have been extensively applied to detect vegetation and monitor vegetation condition over large areas. These indices are generally based on the capability of vegetation to strongly reflect incident electromagnetic signal in the near-infrared (NIR) band compared to the optical bands. In this study we calculated four vegetation indices as candidate predictor variables for the RF classification algorithm to help separate crop and non-crop zones.

NDVI

The Normalized Difference Vegetation Index (NDVI) is commonly used in satellite remote sensing-based vegetation analysis [2,36–38]. It is computed using the red (B4) and near-infrared (B5) bands following Equation (1).

$$NDVI = (B5 - B4) / (B5 + B4) \quad (1)$$

The NDVI can effectively detect growing vegetation [39] but gets quickly saturated in high biomass surfaces. In such conditions, other vegetation indices like EVI (Enhanced Vegetation Index) have been proposed to replace or supplement the NDVI.

EVI

The Enhanced Vegetation Index (EVI), described by Equation (2) provides improved sensitivity to vegetation condition and changes in high biomass areas as compared to the NDVI and also reduces the background effect of soil on vegetation index calculation [40,41]. In addition to the red and near-infrared bands, EVI includes in the calculation the blue band (B2) to correct atmospheric effects of aerosol.

$$EVI = G * ((B5 - B4) / (B5 + C1 * B4 - C2 * B2 + L)) \quad (2)$$

where G is a gain factor; C1 and C2 are the coefficients of the aerosol resistance term, which uses blue band B2 to correct for aerosol influences in the red band B4, and L is the soil-adjustment factor as in SAVI. In this we used the coefficients adopted in the MODIS EVI algorithm, which are L = 1, C1 = 6, C2 = 7.5, and G = 2.5 [42].

SAVI

The soil-adjusted vegetation index (SAVI; Huete, 1988) was developed to compensate for the effects of the soil background in sparsely vegetated areas. Equation (3) is the commonly used expression of SAVI with a soil adjustment factor L. This factor is found to reduce soil noise using the value L = 0.5 for a wide range of vegetation classes [43].

$$SAVI = ((B5 - B4) / (B5 + B4 + L)) * (1 + L) \quad (3)$$

MSAVI

The Modified Soil-Adjusted Vegetation Index (MSAVI; [44]; Equation (4)) was proposed as an improved version of SAVI that minimizes the effect of bare soil [44].

$$MSAVI = \left(2 * B5 + 1 - \sqrt{((2 * B5 + 1)^2 - 8 * (B5 - B4))} \right) / 2 \quad (4)$$

2.3. Random Forest (RF)

We used the random forest (RF) technique as the main classification algorithm in this study. The RF model is an ensemble learning algorithm that can be used to predict both continuous (regression) and categorical (classification) responses. For a classification problem, the response variable is a class which links certain independent values to one of the categories present in the dependent variable [45]. An RF

model comprises an ensemble of decision trees, where each tree constitutes a classifier, which can predict the response-variable using a random sub-sample of the independent variables and observations. Each tree uses a sub-ensemble of training values chosen randomly with replacement (i.e., bootstrap sample). The optimum number of predictors used to split data at each tree's node is $\log(m+1)$, where m is the total number of predictors involved. An ensemble of diverse trees minimizes the effect of bias from individual trees considerably improving the overall predictive accuracy of the model. The final class prediction is chosen by a maximum vote (classification). It has been shown that by increasing the number of trees in the model, the errors of prediction (also known as out-of-bag errors or OOB errors) converge, reducing problems with overfitting [45]. In this study, we used OOB error estimation during the training process to finetune RF model parameters and provide internal cross-validation before independent accuracy assessment.

2.4. Gridding and Accuracy Metrics

A grid of 100 km by 100 km squares was created using ArcMap based on the extent of the available training data. In total, 267 squares (labelled from S1 to S267 in Figure 1) were generated covering the study area. For simplicity, Figure 1 shows the positions and labels of the first and last grid-squares. Satellite image (Landsat) data for each square was classified independently using the RLCM reference data to train and evaluate local RF models.

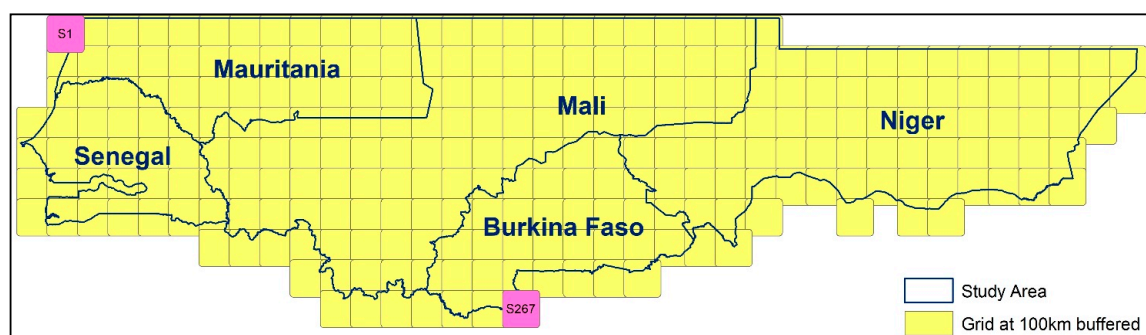


Figure 1. The grid of 100×100 km cells.

Classification accuracy in this study was measured using the following metrics: Quantity disagreement (Q), Allocation disagreement (A), Overall Accuracy (OA), and class-specific measures such as User's Accuracy (UA) for Crop class, as suggested by Pontius and Millones [46]. Appendix A.4 gives more details on disagreements analysis. The new dataset was also validated using detailed local field surveys conducted on agricultural activities at IPR/IFRA (Institut Polytechnique Rural de Formation et de Recherche Appliquée), a higher education institution in Mali.

2.5. Workflow

Figure 2 illustrates the steps employed in developing the cropland extent map, using the GEE platform and Random Forest machine learning approach. Landsat-8 images for each 16-day period were processed for each $\sim 1^\circ$ grid cell. The Tier 1 Landsat-8 image collection was filtered spatially (Sahel grid level) and temporally (years 2013, 2014, and 2015, to match training data) before being filtered for clouds using the pixel_qa information (Table 1). Cloud-free images were used to compute vegetation indices in GEE using custom functions in JavaScript.

In total, twenty bands were exported from GEE as candidate model predictors. Predictors included Landsat 8 surface reflectance bands (B2-B7) and VI averages for growing season (e.g., B2) and dry season (e.g., B2_1), with growing season defined from July–October of each year and November–June considered the dry season. In total, we have twelve (12) surface reflectance and eight (8) vegetation indices (Table 2).

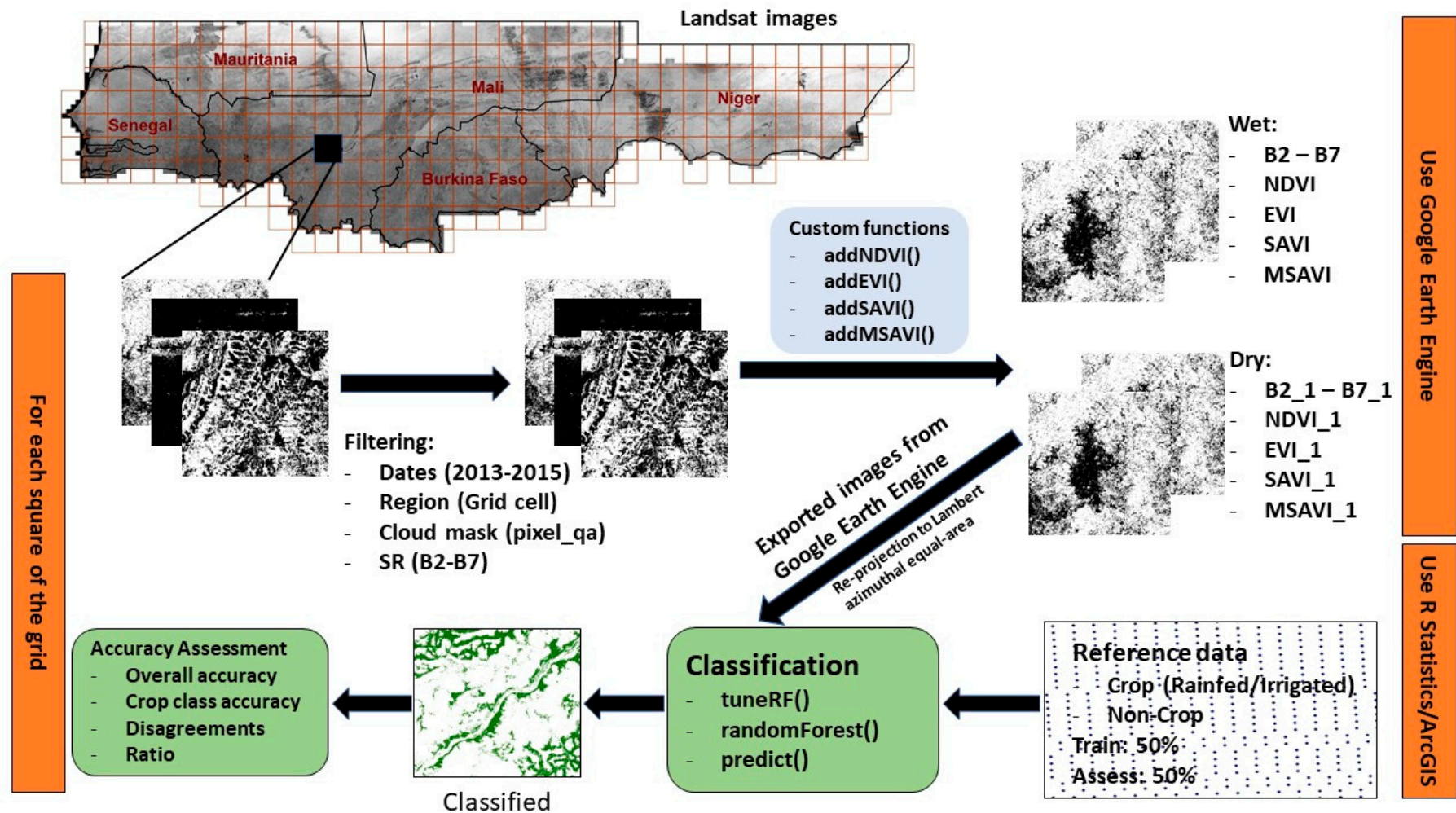


Figure 2. Workflow describing major steps of the cropland dataset development.

Table 2. Predictors used in the Random Forest classification.

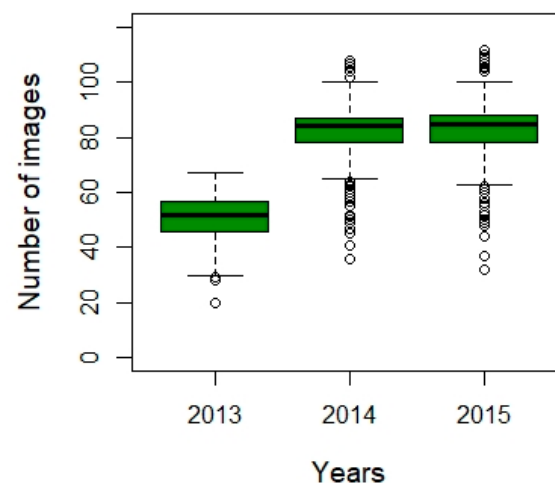
	Wet Period (Growing Period)	Dry Period
Surface Reflectance	B2, B3, B4, B5, B6, B7	B2_1, B3_1, B4_1, B5_1, B6_1, B7_1,
Vegetation Indices	NDVI, EVI, SAVI, MSAVI	NDVI_1, EVI_1, SAVI_1, MSAVI_1

The land-cover training data were reclassified to produce three classes corresponding to rain-fed crops (level “1”), irrigated crops (level “2”) and the non-crop class (level “0”), with levels 1 and 2 combined as needed to make up the “Crop” class. The training samples were derived by sampling 50% of “Crop” and 50% of “Non-crop” classes selected randomly within each grid cell, representing a stratified random sampling approach. R software was used to fit the RF classifiers external to GEE to benefit from greater model-fitting flexibility in R. Optimal fitted RF models were then used for regional predictions. Predictions of crop (rainfed and irrigated) and non-crop classes were based on the best-fit models, optimal parameters, and non-correlated predictors from the tuning process (see Appendix A.1 for more details on optimal RF parameters, and Appendix A.3 for removing correlated variables). Classification outputs were initially assessed at grid-cell level using independent reference samples (i.e., samples not used for training and/or OOB error estimation), then grouped at country and regional levels.

3. Results

3.1. Predictors

The Landsat-8 Tier 1 image collection available on GEE for the study area comprised 6803 image scenes as filtered for the years 2013, 2014, and 2015. Depending on the location and the time, the number of available images changes. This is due, for example, to the degree of cloud coverage in different years and locations. Figure 3 shows that the availability of Landsat images increases in 2014 and 2015, relative to 2013, with the increase related to launch and partial Landsat-8 collection in 2013.



electromagnetic spectrum and low surface vegetation moisture on average in the savanna areas, which would otherwise lower SWIR reflectance. The second highest value occurs at the near infrared band (B5). This band also shows the most pronounced difference between wet and dry means, showing its sensitivity to green vegetation that is mostly present in the wet seasons. All the vegetation indices (Figure 4) show net distinctions between the wet and dry periods, particularly in the range of wet season values.

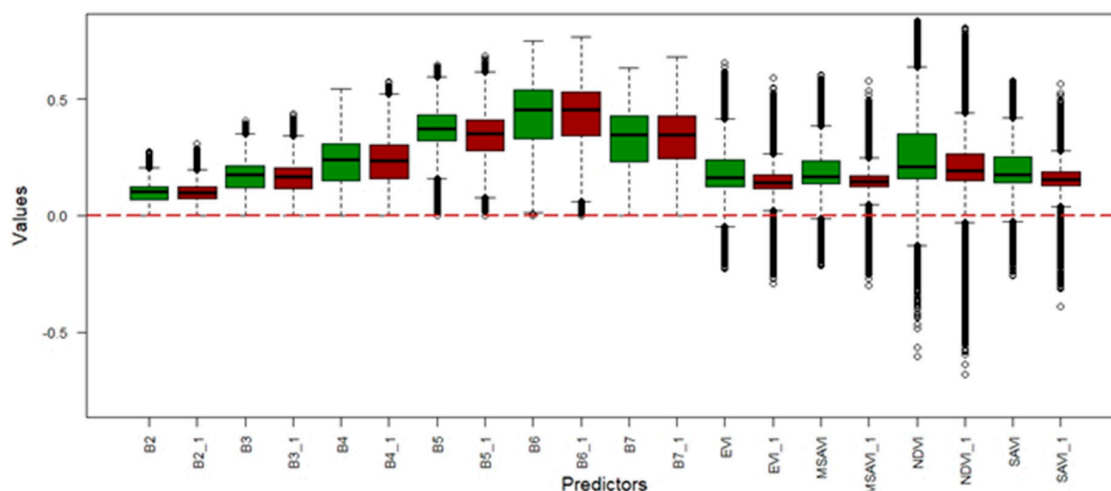


Figure 4. Predictor variables used to train the Random Forest models, showing wet-season averages (green) and dry-season averages (red) for Landsat surface reflectance (SR; bands 2–7) and vegetation index band combinations, averaged across the West African domain (Figure 5). Mean, standard deviation, percentiles, and outliers are as noted for Figure 3.

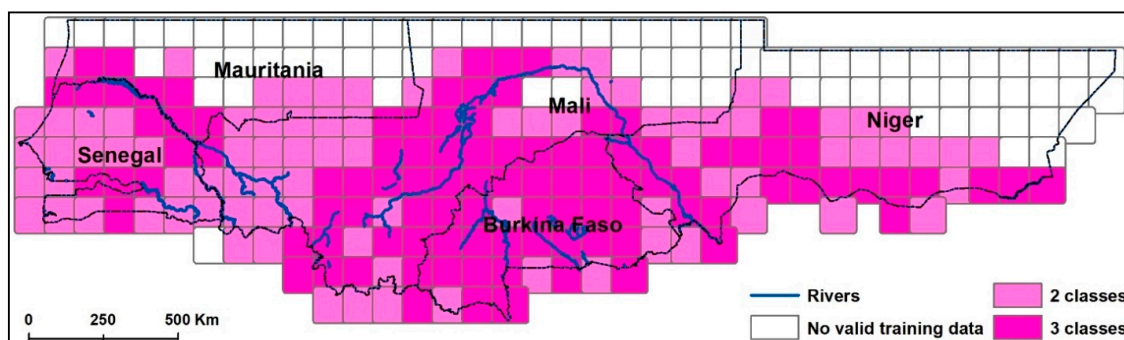


Figure 5. Cells with valid training data (i.e., containing two or more classes (crop and non-crop) after reclassification of the training grid points). A total of 189 cells among the 267 include some cropland in the training data allowing us to run the RF algorithm. 3-classes occur only in those regions with irrigated cropland, mostly associated with the major rivers in the region.

3.2. Reclassified Training Data

Absence of cropland in the training data examined for some grid squares prevented fitting meaningful local models in these regions. These grid squares are therefore assumed to have little or no agriculture (Figure 5). Some cells, particularly in the northern drylands lacked any training data (RF algorithm requires >1 class in the training data). In total, 189 cells (~71% of the study domain) include some amount of cropland. The other 78 cells (white cells in Figure 5) are mainly located in the Northern Sahel and Sahara, where agricultural activities are absent (or occur only intermittently). On average, 2028 reference data on presence of rainfed and irrigated cropland and non-cropland were available in each of the 189 retained grid cells (~1014 for model training and ~1014 for error assessment).

3.3. Accuracy at Grid Level

Results show an average overall accuracy (OA) above 80%, with most 100 km squares having an OA in the range of 75% to 100% (Table 3). Despite the relatively high OA, the reliability of classified product is best measured in terms of the users' accuracy, which quantifies accuracy from the perspective of the user of the classified product. In total, 11% of assessed cells had a user's accuracy of less than 50%; 58% were between 50% and 75%, and the remaining 31% had a user's accuracy above 75% (Table 3). On average, accuracy at grid level was 78.8% and 56.6% for OA and UA, respectively. Please refer to Appendix A.2 of this document for further details on grid level assessment.

Table 3. Summary of overall accuracy and crop class user's accuracy at grid level.

	Overall accuracy			
	No data	0–50	50–75	75–100
Number of cells	99	0	36	132
Average OA	-	-	70.59	87.07
	User's accuracy			
	No data	0–50	50–75	75–100
Number of cells	99	18	98	52
Average UA	-	22.73	64.10	82.95

3.4. Accuracy at Country Level

Assessment at country-scales indicates that the overall accuracy is around 90% for all the countries, except for Burkina Faso where it is slightly lower at ~77% (Figure 6). The country of Mauritania has the highest overall accuracy of 99%, but the accuracy to reliably identify crop class from the user's perspective in Mauritania is only about 71%, which is the lowest among the five countries.

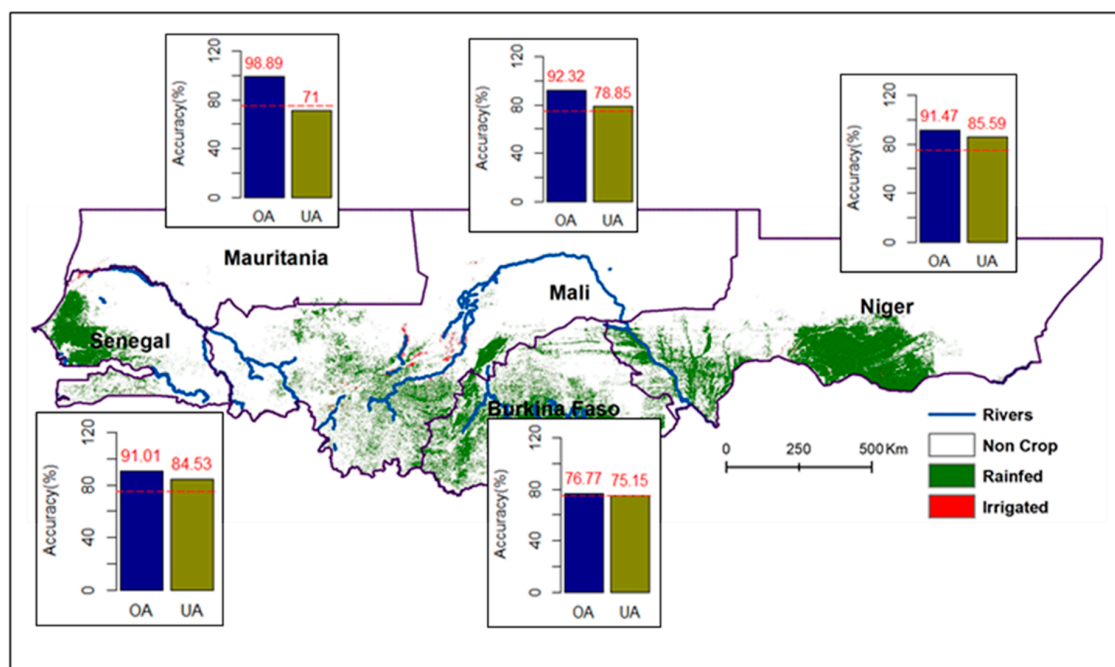


Figure 6. Overall accuracy and user's accuracy by country.

Assuming 75% as targeted value for crop user's accuracy, Mauritania is the only country where the classification performance fails to meet expectations. Highest accuracies occurred in Niger, with 85.6% accuracy, followed by Senegal, with 84.5%. Crop user's accuracy in Mali and Burkina Faso was between 75% and 80%.

In terms of crop area estimation, results show that rainfed agriculture is far more common than irrigated agriculture in the 5 countries, with irrigated cropland occupying only ~2% of the total (Figure 7). Cropland area is greatest in Niger (with $119 \times 10^3 \text{ km}^2$ of cropland, representing 37.6% of the total agricultural area in the five Sahelian countries), followed by Burkina Faso ($91 \times 10^3 \text{ km}^2$; 28.8%), Mali ($67 \times 10^3 \text{ km}^2$; 21.3%), Senegal ($38 \times 10^3 \text{ km}^2$; 12.1%), and finally Mauritania ($0.6 \times 10^3 \text{ km}^2$; 0.21%) where agriculture is confined to the south of the country and the Senegal River Valley (Table 4 and Figure 7).

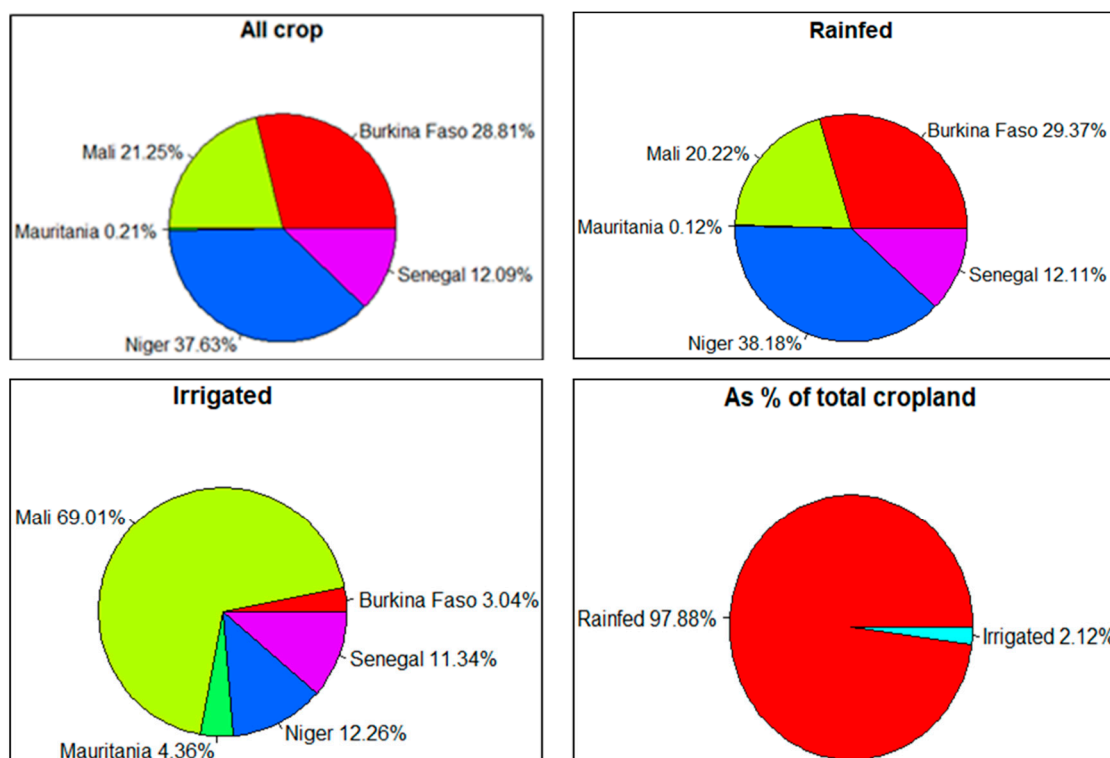


Figure 7. The relative importance (%) of the five Sahelian countries in cropland area across West Africa, showing total cropland area, rainfed and irrigated croplands, and the comparison of rainfed and irrigated cropland area as a fraction of the total.

Table 4. Estimated area classified as cropland in each of five Sahelian countries.

	Burkina Faso	Mali	Mauritania	Niger	Senegal	Total
Rainfed crop (km²)	90,799	62,513	372	118,022	37,434	309,139
Irrigated crop (km²)	203	4615	291	820	758	6688
Total	91,002	67,128	664	118,841	38,192	315,827

Expressed as fraction of the total irrigated area in the five countries, more than the half of the total irrigated areas are in Mali (69.0%), particularly in the “Office du Niger” region, which is one of the oldest and largest irrigation schemes in West Africa [47]. The country of Mauritania, with less than 1% in rainfed agriculture area, has a larger share (4.4%) of the irrigated cropland in the region, more for example than in Burkina Faso with only 3.4% of irrigated cropland (Table 4 and Figure 7).

4. Discussion

The 30 m West African Sahel Cropland map (WASC30) covers five Sahelian countries of West Africa and shows in much more detail than previously available the agricultural zones of West Africa, including the “breadbasket” regions of Niger, Mali, Burkina Faso, and Senegal that are critical to the food security and economies at national and regional scales. We leveraged a distributed and dense sample

dataset on actual land cover [29], with Landsat-8 data, to train locally optimized machine-learning predictors for rainfed and irrigated agriculture using the Google Earth Engine (GEE) platform. Earlier cropland products, covering the Sahelian region, generally combine irrigated and rainfed agriculture into a single cropland class, with accuracies generally less than 70% [15]. The average user's accuracy of the new crop extent map, considering the five countries of Burkina Faso, Mali, Mauritania, Niger, and Senegal (Figure 6), is 79%, which is a considerable improvement relative to the best performing earlier products (GlobeLand30, 69%, and GFSAD30, 64%; [15]). Our accuracy statistic is also influenced by low accuracy in Mauritania, representing less than 1% of cropland area in the region (Figure 7). The user's accuracy for Senegal, Mali, Burkina Faso, and Niger (excluding Mauritania) is 81% for the new WASC30 cropland area map. The low accuracy reported for Mauritania is consistent with our previous findings in a study comparing accuracy of cropland classes in 12 pre-existing landcover products (Samasse et al., 2018). Explanations for this include the particularly small size of farms and the low intensity of agricultural activities in this country. However, it must be noted that, despite having lower accuracy compared to other countries, the cropland estimates for Mauritania in our new WASC30 map are an improvement on the pre-existing products.

4.1. Irrigated Cropland

Based on the estimated crop areas (Table 4), irrigated land represents just 2% of the total cropland area. Thus, a specific accuracy is not reported for this sub-class of "Crop". However, Figure 8 shows clearly the intensive irrigation activities in Senegal and Mauritania adjacent to the Senegal River, in Mali in the "Office du Niger" zone, and in Niger adjacent to the Niger River. Irrigated croplands in the region are generally supported by hydroelectric dams on the major rivers (e.g., Niger, Senegal), providing both electricity and increased agricultural production. For example, the Dama dam in Senegal and the Markala dam in Mali are two operational hydroelectric infrastructures promoting intensive irrigated crop production in the Senegal valley and the Office du Niger zone in Mali, respectively [48,49]. In Mauritania, 44% of the 664 km² mapped as cropland is irrigated.

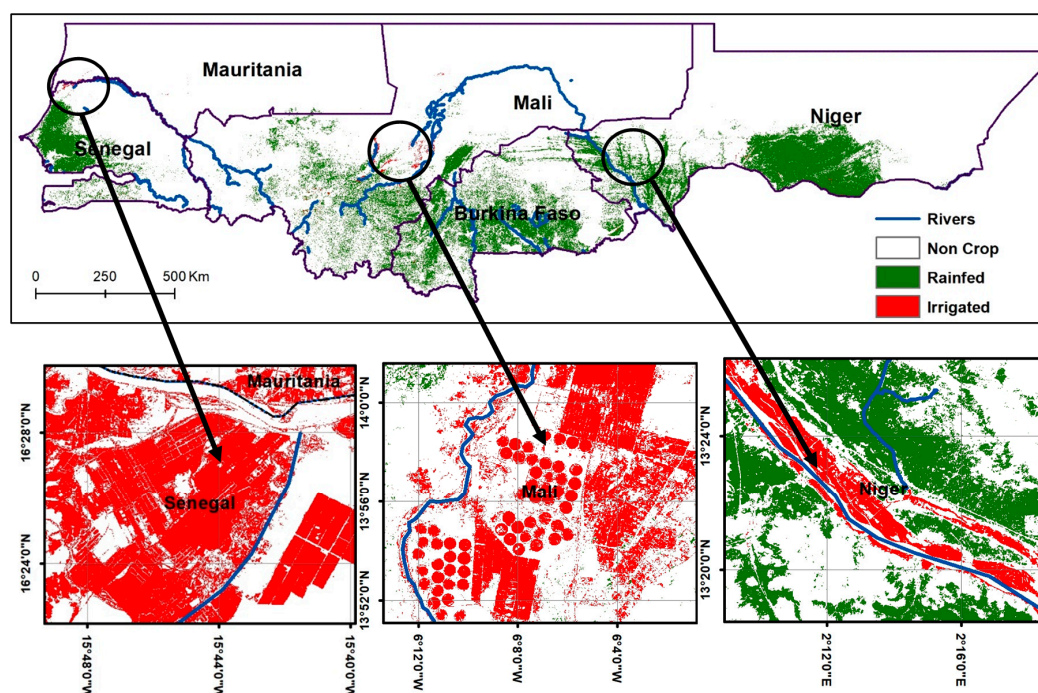


Figure 8. Irrigated cropland adjacent to the Senegal River in South Mauritania and North Senegal, in the Niger River floodplain of Central Mali (with center-pivot irrigation techniques), and adjacent to the Niger River, near Niamey. Rivers are extracted from Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS). Extracted areas are 12 × 12 km.

4.2. Intensive Rainfed Cropland Zones

Analysis of the more dominant rainfed cultivated areas shows several “hot spots” of intensive agricultural activities (Figure 9). For example, the Seno Plain (red circle), east of the Dogon Plateau in Mali, has been devoted to intensive agricultural activities since the 1930s [50]. Recent studies using Earth Observation data have reported cropland expansion in this region driven by the need to feed a rapidly increasing population with accelerated expansion between 2000 and 2013 facilitated by modern technology [32]. Rapid population growth and conducive soils, with development of processing infrastructure, have also contributed to the high density of rainfed cropland in south-eastern Niger (blue circle). This area, known as the Tarka Plain and Goulbi Agricultural Zone, in the Maradi-Zinder region of Niger, is considered the most important agricultural zone of Niger (CILSS, 2016). It is an area of enormous agricultural potential, mainly in rainfed cropland [51]. Cereal (Millet, Maize, Sorghum, and Rice) cultivation is practiced, with more advanced systems in the Tarka plain, where the rural population density is particularly high. Because of the anthropogenic pressure, it is common to see an integrated system where agriculture, livestock, and forests share the same space [52]. Similarly, the West-Central Agricultural Zone in Senegal (black circle), known as the Peanut Basin (Bassin Arachidier) for the suitability of dominant soils to grow peanuts, is also characterized by high rural populations, with rainfed agriculture focused on cultivation of peanut, millet, sorghum, and beans [53].

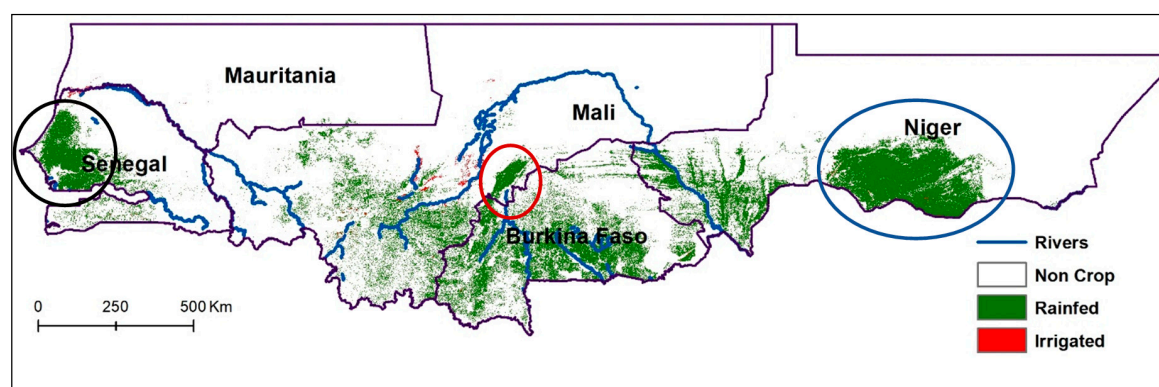


Figure 9. The West African Sahel cropland map (WASC30) with hotspots of intensive rainfed cropland in Senegal, Mali and Niger.

4.3. Cropland Distribution Relative to Climate and Climate Zones

For the purpose of this work, we divided the study area based on the annual precipitation, following a steep gradient of decreasing rainfall from south to north. Figures 10 and 11 show the distribution of both rainfed and irrigated cropland as located in the West African Sahel Cropland map (WASC30) under 100 mm rainfall bins. Mean Annual Precipitation (MAP) is derived from eleven years (2005–2015) of CHIRPS (Climate Hazards Group InfraRed Precipitation with Station) data retrieved from Google Earth Engine. Results show that rainfed cropland area generally increases with MAP between 200 and 1000 mm MAP, reflecting water limitations to agricultural activities in the arid zones (200–400 mm) and more suitable conditions in the South Sahel and Soudan (400–1000 mm). Above 1000 mm MAP, rainfed cropland proportion declines (Figure 11), in part due to shift to forest production and in part since wetter forested zones may have soils unsuitable for agriculture. However, irrigated cropland area is largely decoupled from MAP, being clustered around the flood plains of the perennial rivers in West Africa.

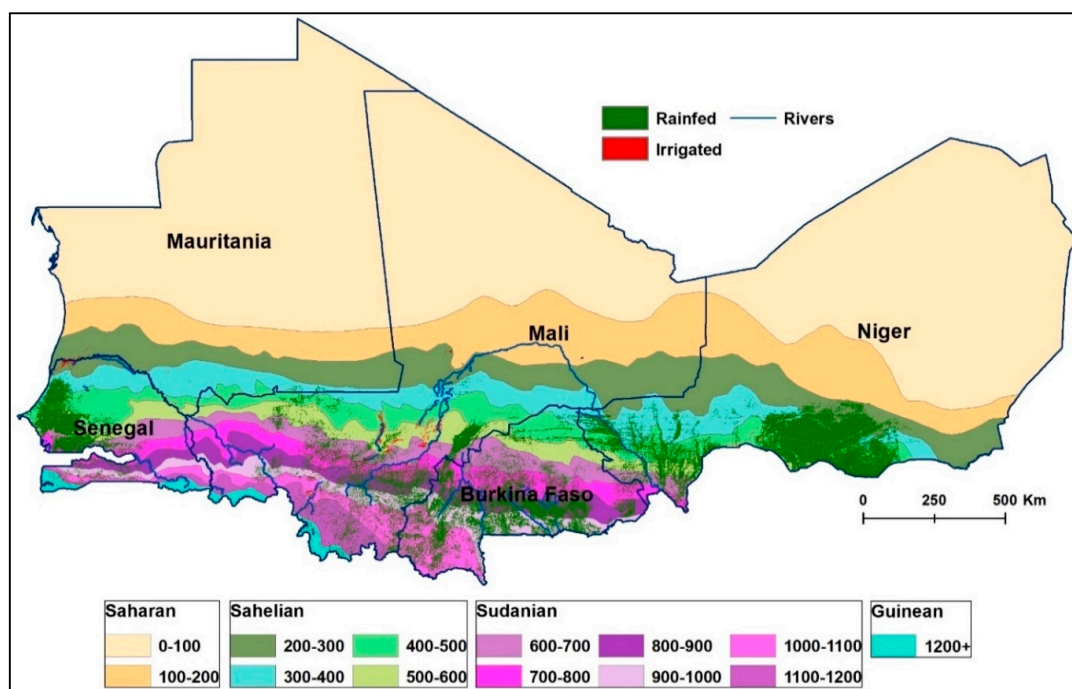


Figure 10. Cropland extent across gradient of mean annual precipitation. Precipitation is an average of 11 years (2005–2015). CHIRPS (Climate Hazards Group InfraRed Precipitation with Station) data.

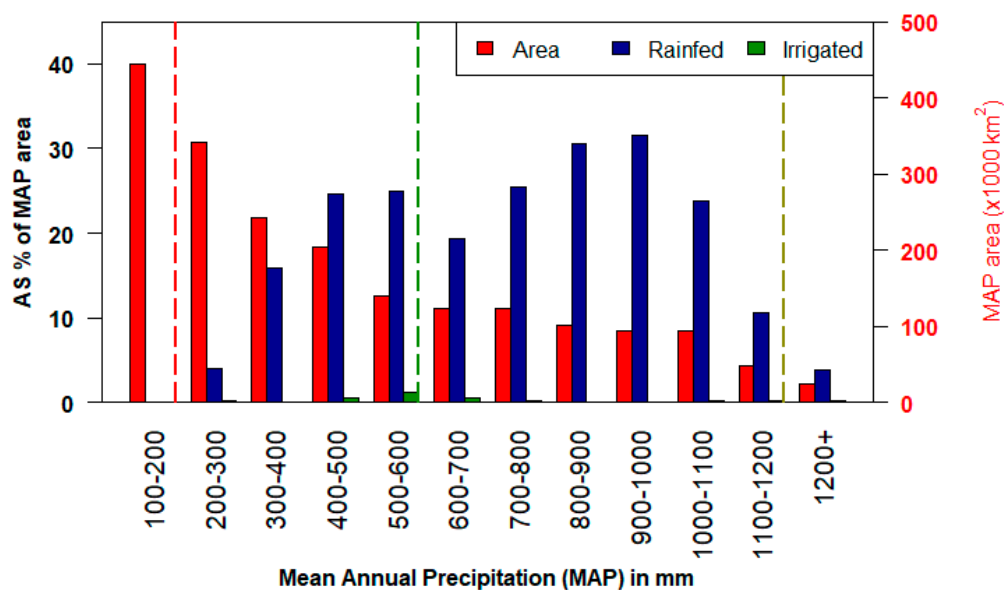


Figure 11. Rainfed and irrigated cropland as percentage of the total area in MAE interval (e.g., The total area of 800–900 mm zone is about $100 \times 10^3 \text{ km}^2$, and the fraction of this area occupied by rainfed cropland is about 30%).

The Saharan desert region (MAP < 200 mm/y; Figure 10) constitutes about 61% of the total study area (Table 5). Significant part of northern Mauritania, Mali, and Niger fall in this region. It is generally characterized by an arid climate with high average temperatures, a very low relative humidity, and rare and highly irregular precipitation, making difficult for crops to grow. However, irrigated farming may be present in some areas using appropriate irrigation technologies [54,55], mainly for small scale production of vegetables. Figure 11 shows our results illustrating the very low to non-existent agricultural activities in the Sahara.

Table 5. Proportion of area and cropland by climatic zones in our study domain including Senegal, Mauritania, Mali, Burkina Faso, and Niger.

	Saharan	Sahelian	Sudanian	Guinean
% Area	61.02	23.50	14.85	0.63
% Rainfed	0.00	14.79	24.54	3.78
% Irrigated	0.01	0.48	0.26	0.17
% All Cropland	0.01	15.27	24.80	3.95

The second largest zone is the Sahelian (200–600 mm), occupying 23% of the total study area. Rainfed cropland intensity increases with annual rainfall (Figure 11). Compared to the other climatic zones, the Sahel has the highest proportion of irrigated cropland, as irrigation activities along both Niger and Senegal rivers occur mainly in this climatic zone (Figure 10). This irrigation proportion is, however, less than 1% against about 15% for rainfed agriculture (Table 5).

Among the four climatic zones (Figure 10), rainfed agriculture activities are most intensive in the Sudanian zone (600–1200 mm). Representing 15% of the study area, the Sudanian is the third largest climatic zone, after the Sahara and the Sahel. About 24% of this climatic zone is occupied by rainfed cropland (Table 5). It covers major cereal production zones in Mali and Burkina Faso and southern parts of the Peanut Basin in Senegal (Figure 10). The precipitation range is also suitable for cash crops (e.g., cotton), root crops, and mixed cereal-root system (e.g., cassava, yam, sweet potato, particularly in Southern Mali). Irrigation is not common in the Sudanian zone, largely due to the low occurrence of main rivers in the region.

At more than 1200 mm MAP, the Guinean zone covers little of the total area of the study domain (less than 1%). In this region, some 4% is occupied by rainfed cropland, dominated by root crop cultivation (yams, sweet potatoes, cassava) [56].

4.4. Fallows in WASC30

The visual interpretation approach adopted for the 2 km RLCM dataset [29,31] and used as training or reference information in this study classified large and long-term fallows as savanna. However, in the more intensive rainfed cropland regions with reduced fallow periods, small areas of fallow were generally classified as active agricultural land use. Overall, therefore, we consider that the WASC30 represents active agriculture, inclusive of short-term fallow fields but exclusive of longer-term fallow (or abandoned) areas that have not been actively cropped in recent years. That makes the final cropland class a reliable reference for developing active cropland extent.

4.5. Validation Using Local Scale Data

At local scale, our new cropland dataset has been assessed using recent (2012) GPS field surveys mapping land use and land cover at an Agricultural College (IPR/IFRA (Institut Polytechnique Rural de Formation et de Recherche Appliquée)) in the town of Koulikoro, just north of Bamako. Data was collected in collaboration with Laval University (Quebec, Canada) as part of the PACM research project (Des arbres et des champs contre la pauvreté au Mali). IPR/IFRA is a higher education institution in Mali managing an area of about 380 ha, including experimental farms and other lands for cereals and tree crop production. Figure 12 shows that the WASC30 captures the distribution of cultivated areas at IPR/IFRA with an area ratio of 195 ha/199 ha = 98% (i.e., the WASC30 product underestimates cropland area at this field station by 2%). This slight difference could be attributed to the small size of some sparse experimental plots making difficult their detection in the 30 m Landsat data. No irrigated pixels were detected at IPR/IFRA, which is consistent with the absence of irrigation trials at the site.

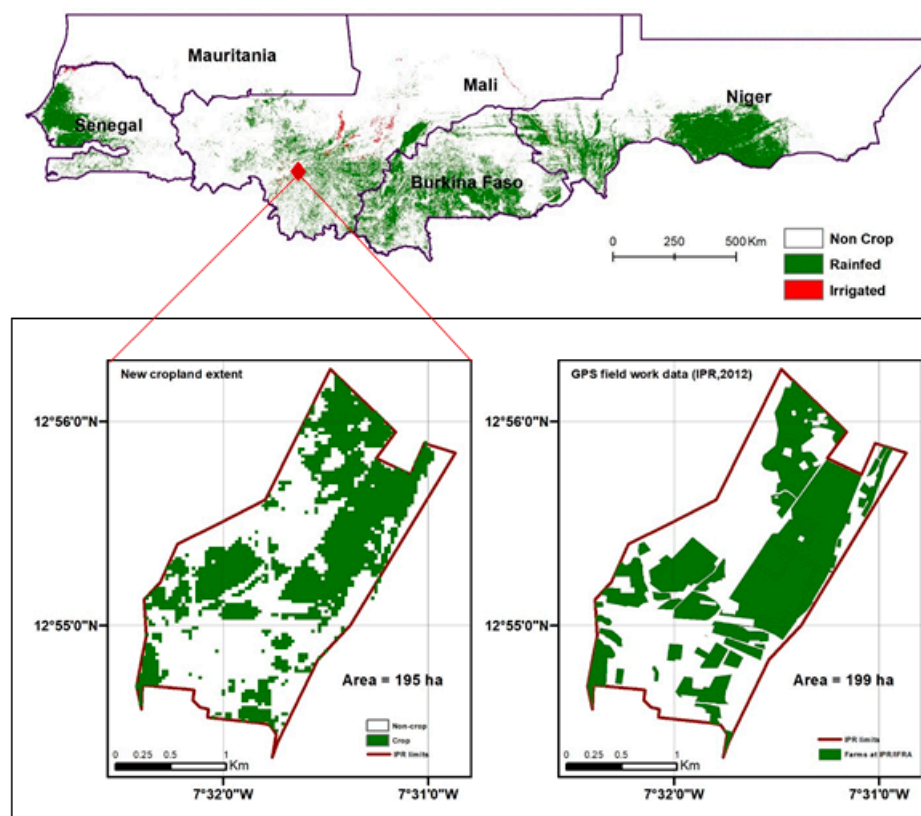


Figure 12. Assessment of the new cropland map (WASC30) using field surveys at IPR/IFRA (Institut Polytechnique Rural de Formation et de Recherche Appliquée) field station in Mali.

5. Conclusions

In this study, the Random Forest ensemble learning method has been applied to individual 100 km grid cells to develop a 30 m Landsat-derived active cropland dataset across five Sahelian countries with unprecedented details and higher accuracy as compared to existing land cover products. The developed dataset has an overall accuracy of 90.1% and a cropland class (rainfed and irrigated) user's accuracy of 79%.

Information derived from the new dataset reveals the total cropland area in West African Sahel to be $316 \times 10^3 \text{ km}^2$ with $7 \times 10^3 \text{ km}^2$ irrigated and $309 \times 10^3 \text{ km}^2$ rainfed. This confirms that agriculture in Sahelian West Africa is almost entirely rainfed. The Sudanian zone (600–1200 mm) comprises most of the rainfed cultivated areas, while the Sahelian areas in proximity to main rivers present the highest proportion of irrigated land. Results also show that these irrigation activities in the region remain not well developed, comprising only about 2% of the total cropland area, despite the tremendous potential offered by, for example, the Senegal and Niger rivers. This may be due to the lack of well-developed infrastructure for irrigation and high investment costs to manage water and make it available where it is most needed. More efforts in developing irrigated land in Sahel region would expand farmers' production opportunities by reducing risks linked to climate fluctuations.

This study benefitted from the large and regularly distributed RLCM training dataset that allowed us to fit locally optimized random forest models in each of 189 grid cells (each $100 \times 100 \text{ km}$) across the five-country study domain. This allowed us to minimize the effects of soil, topographic, and climatic differences that would increase errors in models fit at coarser regional and continental scales, thus improving overall accuracy of the final WASC30 product.

Geospatial data in general, and Landsat time series in particular, provide a critical source of information for the important task of producing accurate statistics on cultivated areas, particularly in

developing countries where timely accurate georeferenced agricultural data are sometimes missing. The new cropland dataset will contribute to filling this void in West Africa Sahel.

Data Availability: The West African Sahel Cropland Dataset (WASC30) is currently available for visualization as a Google Earth Engine Application (<https://savannalabnmsu.users.earthengine.app/view/wa-cropmap-30m>). Further inquiries on data availability including download options should be addressed to the corresponding author.

Author Contributions: K.S. and N.P.H. conceived the analysis. K.S. and J.Y.A. carried out the analysis and K.S. wrote the manuscript. Y.D. helped with IPR/IFRA data for validation. All authors contributed to writing and editing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the United States Agency for International Development (USAID), as part of the Feed the Future initiative, under the CGIAR Fund, award number BFS-G-11-00002, and the predecessor fund, the Food Security and Crisis Mitigation II grant, award number EEM-G-00-04-00013 (USAID) via funding to the Michigan State University Borlaug Higher Education for Agricultural Research and Development (BHEARD) Graduate Research Fellowship Program. N.P.H. and J.Y.A. were supported, in part, by the NASA SERVIR West Africa Program (Grant # NNX16AN30G).

Acknowledgments: Our particular thanks to Dr. Gray Tappan and colleagues at USGS EROS Data Center who made available the West African Land Cover time series database (<https://pubs.er.usgs.gov/publication/fs20173004>) that was invaluable as training and validation for this analysis.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A.

Appendix A.1. Tuning RF Major Parameters

Separate RF models were developed for each of the 189 cropland cells. Correlated predictors were removed following the example illustrated by Figure A4 before fitting models and tuned using the “tuneRF” function in R software. This tuning function helps determine the best number of variables available for splitting at each tree node (mTry) for a number of trees (nTree) based on the minimum values of the Out-Of-Bag (OOB) errors. The chosen values of nTree to run “tuneRF” included 500, 1000, 1500, and 2000. Occurrences of nTree corresponding to the optimal mTry are reported in Figure A1A. It appears that RF models (classifiers) show better performance in 44% of the cells for nTree = 500 and 15% of the cells for nTree = 2000. Between these two limits, 24% and 17% of the cells have shown minimal errors of OOB at nTree = 1000 and nTree = 1500, respectively. Frequency distribution of resulting optimum mTry values from the tuning process is shown on Figure A1B. Numbers 2, 1, 4, 8, 16, and 3 have been used as mTry values to fit the best models for the classification.

The Random Forest (RF) is widely accepted as an efficient ensemble approach for land cover classification using remotely sensed data. It handles well imbalanced data, missing values, and outliers [57]. However, tuning RF two major parameters (number of trees: nTree, number of variables available for splitting at each tree node: mTry) to get optimum values may be time and resource consuming, even in parallel processing environments like Google Earth Engine. In this study, we selected nTree in {500, 1000, 1500, 2000} for reduced computational time while ensuring sufficient trees for model convergence (Breiman, 2001). The best mTry for most of the grid cells has been achieved with nTree = 500, others for nTree = 2000, which are the limits of tuned nTree values (Figure A1A). Using a larger range of nTree values (e.g., including values below 500 and above 2000) could probably result in better mTry, yielding higher classification performance for the final cropland product.

Appendix A.2. Accuracy at Grid Level

Results on the accuracy assessment do not show all the 189 trained and classified squares (100 km by 100 km grid unit). Twenty-one (21) of them have been entirely classified as Non-crop. They are considered as NoData for the assessment. The general trend appearing on Figure A2 is that in average classified squares have an overall accuracy above 80%. For most squares, overall accuracy falls in the range of 75% to 100%. The country of Burkina Faso has the maximum of units with overall accuracy within 50%–75%, and none in the grid has been classified with a correct proportion of less than 50%.

This relatively high overall accuracy is contrasted by the crop class specific accuracy. Figure A3 gives insight into the “Crop” class user’s accuracy at grid level. In total, 11% of assessed cells have a user’s accuracy less than 50%, 58% of them have theirs between 50% and 75%, and the remaining 31% have a user’s accuracy above 75% (Table 3).

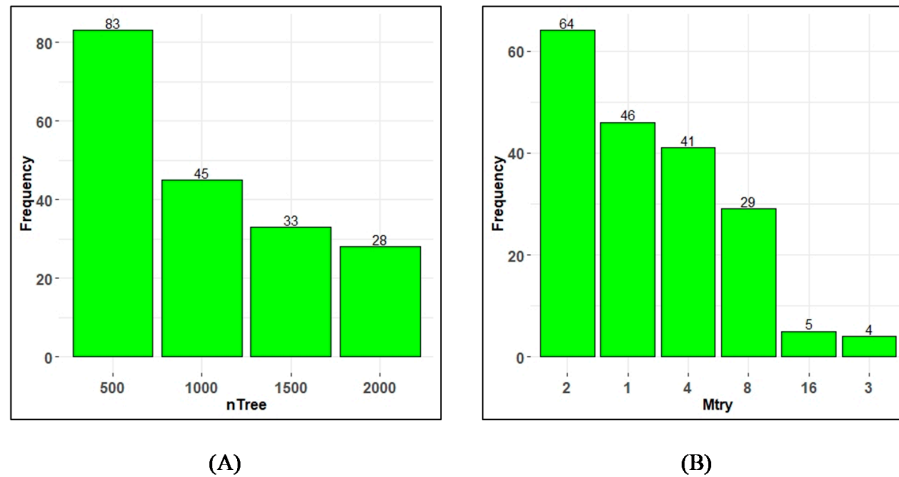


Figure A1. Occurrences of the number of tree (nTree) corresponding to minimum out-of-bag errors (A) and the obtained number of variables available for splitting at each tree node, mTry (B).

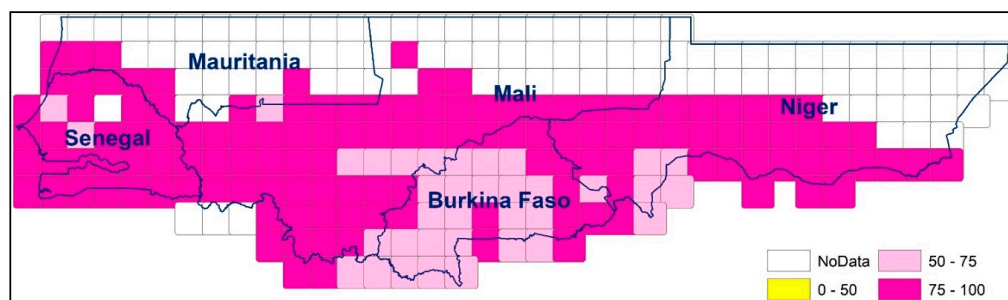


Figure A2. Overall accuracy (OA) at grid level.

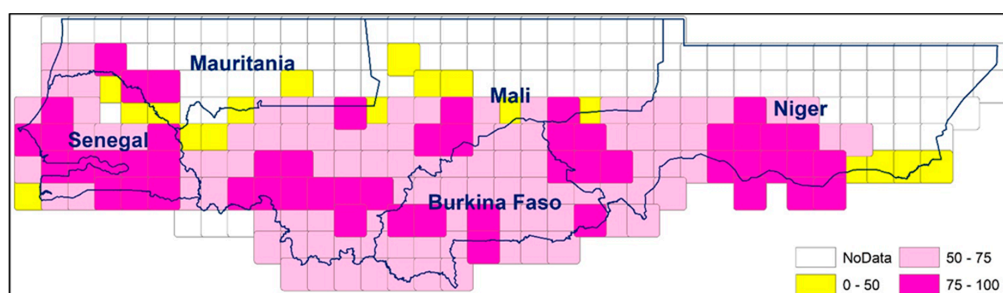


Figure A3. Crop User's accuracy at grid level.

Appendix A.3. Correlated Variables/Predictors

An example of removal of highly correlated variables, based on a correlation coefficient > 0.99 is shown in Figure A4. In these cases, we anticipate no additional useful explanatory information is available by including both variables. Predictor EVI is similar to SAVI and MSAVI, thus only MSAVI was maintained to develop the model by the algorithm. Similarly, predictors EVI_1 and SAVI_1 bring the same information as MSAVI_1; they can then be removed, reducing computational time in classifying this grid cell. Since RF is generally robust to correlation, correlations < 0.99 were permitted.

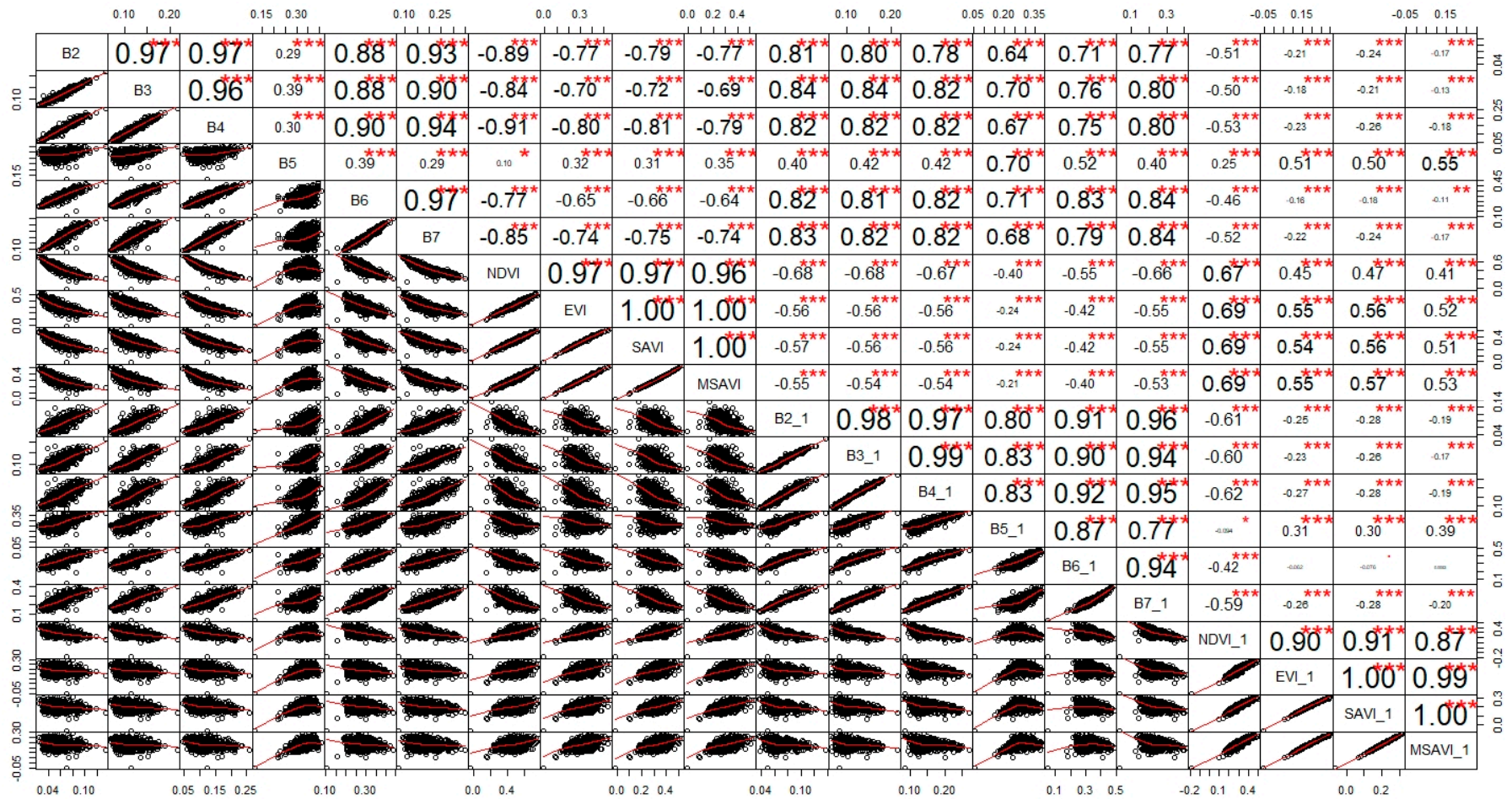


Figure A4. Correlated variables.

Appendix A.4. Disagreements Analysis

Error of classification, expressed as total disagreement, can be divided into two components, which are the quantity disagreement and allocation disagreement [46]. Quantity disagreement can be interpreted as the differences in the areas allocated to the classes in the reference data and the classified map, and allocation disagreement is related to the misallocation of classified pixels for the same level of quantity agreement [58]. The overall accuracy is the complement of the total disagreement (100%—total disagreement). Figure A5 illustrates results of the developed cropland map considering these two categories of disagreement at country level. Overall, the highest total disagreement is less than 25%. That means for the 5 countries, overall accuracy is greater than 75%. It also appears that quantity disagreement is more important in Mali, Mauritania, and Niger. Thus, this measure is the major contributor to the map's total disagreement in these countries. The opposite is true for Burkina Faso and Senegal where spatial mismatch of pixels dominates the disagreement. Considering 10% as threshold of disagreement significance, Burkina Faso is the only country exceeding this level.

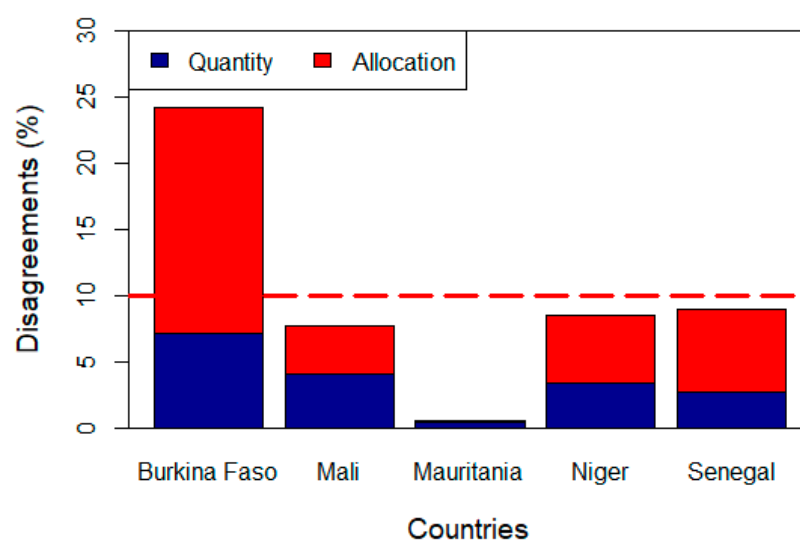


Figure A5. Quantity disagreement (quantity = how much cropland) and Allocation disagreement (allocation = where the cropland is) by country.

References

1. Latham, J. FAO Land Cover Mapping Initiatives. In Proceedings of the North American Land Cover Summit, Washington, DC, USA, 20–22 September 2006; Environment and Natural Resources Service of the Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 2009; pp. 75–95.
2. Thenkabail, P.; Lyon, J.G.; Turrall, H.; Biradar, C. *Remote Sensing of Global Croplands for Food Security*; CRC Press: Boca Raton, FL, USA, 2009.
3. Hollinger, F.; Staatz, J.M. *Agricultural Growth in West Africa: Market and Policy Drivers*; FAO: Rome, Italy; African Development Bank: Tunis, Tunisia; ECOWAS: Abuja, Nigeria, 2015.
4. Atzberger, C. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sens.* **2013**, *5*, 949–981. [[CrossRef](#)]
5. Jones, H.G.; Vaughan, R.A. *Remote Sensing of Vegetation: Principles, Techniques, and Applications*; Oxford University Press: Oxford, UK, 2010.
6. 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*, 232. [[CrossRef](#)]
7. Pérez-Hoyos, A.; Rembold, F.; Kerdiles, H.; Gallego, J. Comparison of global land cover datasets for cropland monitoring. *Remote Sens.* **2017**, *9*, 1118. [[CrossRef](#)]
8. Xiong, J.; Thenkabail, P.S.; Tilton, J.C.; Gumma, M.K.; Teluguntla, P.; Oliphant, A.; Congalton, R.G.; Yadav, K.; Gorelick, N. 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 Sens.* **2017**, *9*, 1065. [CrossRef]
9. Burke, M.; Lobell, D.B. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2189–2194. [CrossRef]
 10. Hong, C.; Jin, X.; Ren, J.; Gu, Z.; Zhou, Y. Satellite data indicates multidimensional variation of agricultural production in land consolidation area. *Sci. Total Environ.* **2019**, *653*, 735–747. [CrossRef]
 11. Löw, F.; Biradar, C.; Fliemann, E.; Lamers, J.P.A.; Conrad, C. Assessing gaps in irrigated agricultural productivity through satellite earth observations—A case study of the Fergana Valley, Central Asia. *Int. J. Appl. Earth Obs. Geoinform.* **2017**, *59*, 118–134. [CrossRef]
 12. Rembold, F.; Meroni, M.; Urbano, F.; Csak, G.; Kerdiles, H.; Perez-Hoyos, A.; Lemoine, G.; Leo, O.; Negre, T. ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. *Agric. Syst.* **2019**, *168*, 247–257. [CrossRef]
 13. Kobayashi, N.; Tani, H.; Wang, X.; Sonobe, R. Crop classification using spectral indices derived from Sentinel-2A imagery. *J. Inform. Telecommun.* **2019**, 1–24. [CrossRef]
 14. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sensors* **2017**. [CrossRef]
 15. Samasse, K.; Hanan, N.P.; Tappan, G.; Diallo, Y. Assessing cropland area in West Africa for agricultural yield analysis. *Remote Sens.* **2018**, *10*, 1785. [CrossRef]
 16. Arino, O.; Bicheron, P.; Achard, F.; Latham, J.; Witt, R.; Weber, J.L. The most detailed portrait of Earth. *Eur. Space Agency* **2008**, *136*, 25–31.
 17. Bartholome, 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. [CrossRef]
 18. Bontemps, S.; Boettcher, M.; Brockmann, C.; Kirches, G.; Lamarche, C.; Radoux, J.; Santoro, M.; Van Bogaert, E.; Wegmüller, U.; Herold, M.; et al. Multi-year global land cover mapping at 300 M and characterization for climate modelling: Achievements of the land cover component of the ESA climate change initiative. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 323–328. [CrossRef]
 19. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [CrossRef]
 20. 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. [CrossRef]
 21. Fritz, S.; You, L.; Bun, A.; See, L.; McCallum, I.; Schill, C.; Perger, C.; Liu, J.; Hansen, M.; Obersteiner, M. Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophys. Res. Lett.* **2011**, *38*. [CrossRef]
 22. 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.
 23. Tong, X.; Brandt, M.; Hiernaux, P.; Herrmann, S.; Rasmussen, L.V.; Rasmussen, K.; Tian, F.; Tagesson, T.; Zhang, W.; Fensholt, R. The forgotten land use class: Mapping of fallow fields across the Sahel using Sentinel-2. *Remote Sens. Environ.* **2020**, *239*, 111598. [CrossRef]
 24. Buchhorn, M.; Smets, B.; Bertels, L.; Lesiv, M.; Tsendbazar, N.E.; Herold, M.; Fritz, S. Copernicus Global Land Service: Land Cover 100 m: Epoch 2018: Africa Demo. Available online: <https://zenodo.org/record/3518087#.XqgPLmgzZPY> (accessed on 20 February 2020).
 25. Li, L.; Tsendbazar, N.; Herold, M.; Lesiv, M. Copernicus Global Land Operations “Vegetation and Energy”: Moderate Dynamic Land Cover Change Maps, Africa 2015–2018. Available online: https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_PUM_LCC100m-V2.1_I3.10.pdf (accessed on 25 February 2020).
 26. 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**, *202*, 18–27. [CrossRef]
 27. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [CrossRef]

28. Kumar, L.; Mutanga, O. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sens.* **2018**, *10*, 1509. [\[CrossRef\]](#)
29. Tappan, G.G.; Cushing, W.M.; Cotillon, S.E.; Mathis, M.L.; Hutchinson, J.A.; Dalsted, K. *West Africa Land Use Land Cover Time Series*; U.S. Geological Survey: Sioux Falls, SD, USA, 2016.
30. Cotillon, S.E. *West Africa Land Use and Land Cover Time Series*; Fact Sheet 2017–3004; U.S. Geological Survey: Sioux Falls, SD, USA, 2017. [\[CrossRef\]](#)
31. Cotillon, S.E.; Mathis, M.L. *Mapping Land Cover Through Time with the Rapid Land Cover Mapper—Documentation and User Manual*; Open File Report 2017–1012; U.S. Geological Survey: Sioux Falls, SD, USA, 2017; p. 23. [\[CrossRef\]](#)
32. CILSS. *Landscapes of West Africa—A Window on a Changing World*; U.S. Geological Survey EROS: Garretson, SD, USA, 2016.
33. Mardani, M.; Mardani, H.; De Simone, L.; Varas, S.; Kita, N.; Saito, T. Integration of Machine Learning and Open Access Geospatial Data for Land Cover Mapping. *Remote Sens.* **2019**, *11*, 1907. [\[CrossRef\]](#)
34. Johnson, D.M. Using the Landsat archive to map crop cover history across the United States. *Remote Sens. Environ.* **2019**, *232*, 111286. [\[CrossRef\]](#)
35. Azzari, G.; Lobell, D.B. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sens. Environ.* **2017**, *202*, 64–74. [\[CrossRef\]](#)
36. Fensholt, R.; Rasmussen, K.; Nielsen, T.T.; Mbow, C. Evaluation of earth observation based long term vegetation trends—Intercomparing NDVI time series trend analysis consistency of Sahel from AVHRR GIMMS, Terra MODIS and SPOT VGT data. *Remote Sens. Environ.* **2009**, *113*, 1886–1898. [\[CrossRef\]](#)
37. Olsson, L.; Eklundh, L.; Ardö, J. A recent greening of the Sahel—Trends, patterns and potential causes. *J. Arid Environ.* **2005**, *63*, 556–566. [\[CrossRef\]](#)
38. Vintrou, E.; Desbrosse, A.; Bégué, A.; Traoré, S.; Baron, C.; Seen, D.L. 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. Geoinform.* **2012**, *14*, 83–93. [\[CrossRef\]](#)
39. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [\[CrossRef\]](#)
40. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [\[CrossRef\]](#)
41. Miura, T.; Huete, A.R.; Yoshioka, H.; Holben, B.N. An error and sensitivity analysis of atmospheric resistant vegetation indices derived from dark target-based atmospheric correction. *Remote Sens. Environ.* **2001**, *78*, 284–298. [\[CrossRef\]](#)
42. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [\[CrossRef\]](#)
43. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [\[CrossRef\]](#)
44. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [\[CrossRef\]](#)
45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
46. Pontius, R.G.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [\[CrossRef\]](#)
47. Zwart, S.J.; Leclert, L.M.C. A remote sensing-based irrigation performance assessment: A case study of the Office du Niger in Mali. *Irrig. Sci.* **2010**, *28*, 371–385. [\[CrossRef\]](#)
48. van der Wijngaart, R.; Helming, J.; Jacobs, C.; Delvaux, P.A.G.; Hoek, S.; Gomez y Paloma, S. *Irrigation and Irrigated Agriculture Potential in the Sahel: The Case of the Niger River Basin*; JRC Technical Report; Publications Office of the European Union: Luxembourg, 2019.
49. Woodhouse, P.; Ganho, A.S. Is Water the Hidden Agenda of Agricultural Land Acquisition in Sub-Saharan Africa. In Proceedings of the International Conference on Global Land Grabbing, Sussex, UK, 6–8 April 2018; Land Deals Politics Initiative: The Hague, The Netherlands, 2011; pp. 1–19.
50. Thibaud, B. Le pays dogon au Mali: De l'enclavement à l'ouverture ? *Espac. Popul. Soc.* **2005**, *1*, 45–56. [\[CrossRef\]](#)
51. Issoufou, W.S.; Mahamane, A.; Ousseini, I. La Surveillance Ecologique et Environnementale au Niger: Un instrument d'aide à la décision. *Options Méditerran. Sér. B. Etudes Rech.* **2012**, *68*, 219–230.

52. RNCA-NIGER. Le Zonage Agro-Ecologique du NIGER. 2019. Available online: https://reca-niger.org/IMG/pdf/Le_zonage_agroecologique_du_Niger_Extraits.pdf (accessed on 15 January 2020).
53. FALL, C.A. État des Ressources Phytogénétiques pour l’Alimentation et l’Agriculture dans le Monde: Contribution du Sénégal au Second Rapport. 2009. Available online: <http://www.fao.org/pgrfa-gpa-archive/sen/docs/senegal2.pdf> (accessed on 15 January 2020).
54. Bouzidi, B. Viability of solar or wind for water pumping systems in the Algerian Sahara regions—Case study Adrar. *Renew. Sustain. Energy Rev.* **2011**, *15*, 4436–4442. [[CrossRef](#)]
55. Hamidat, A.; Benyoucef, B.; Hartani, T. Small-scale irrigation with photovoltaic water pumping system in Sahara regions. *Renew. Energy* **2003**, *28*, 1081–1096. [[CrossRef](#)]
56. Sidibe, A. *L’Etat des Ressources Phytogénétiques Pour l’Alimentation et l’Agriculture au Mali—2007*; Deuxième Rapport National; FAO: Rome, Italy, 2007.
57. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [[CrossRef](#)]
58. Richards, J.A.; Jia, X. *Remote Sensing Digital Image Analysis*; Springer Nature Switzerland AG: Basel, Switzerland, 1999; Volume 3.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).