

September, 2017

**NASA Making Earth System Data Records for Use in
Research Environments (MEaSUREs) Global Food
Security-support Analysis Data (GFSAD) @ 30-m for
South America: Cropland Extent Product
(GFSAD30SACE)**

Algorithm Theoretical Basis Document (ATBD)

USGS EROS
Sioux Falls, South Dakota

Document History

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

Contents

Document History	2
I. Members of the team	4
II. Historical Context and Background Information	5
III. Rationale for Development of the Algorithms.....	6
IV. Algorithm Description	7
a. Input data.....	8
i. Region Definition	8
ii. Reference Samples.....	8
iii. Satellite Imagery: Landsat data	8
b. Theoretical description.....	9
i. Definition of Croplands	9
ii. Algorithm.....	10
c. Practical description	10
i. Random Forest (RF) Algorithm.....	10
ii. Programming and codes	12
iii. Results	12
iv. Cropland areas of South America.....	16
V. Calibration Needs/Validation Activities	17
VI. Constraints and Limitations.....	18
VII. Publications	19
i. Peer-reviewed publications within GFSAD project.....	19
ii. Other relevant past publications prior to GFSAD project.....	21
iii. Books and Book Chapters	22
VIII. Acknowledgments.....	24
IX. Contact Information.....	24
X. Citations	25
XI. References.....	25

I. Members of the team

This Global Food Security-support Analysis Data 30-m (GFSAD30) Cropland Extent Product of South America (GFSAD30SACE) was produced by the following team members. Their specific roles are mentioned below.

Ms. Ying Zhong, Scientist, Environmental Systems Research Institute (ESRI), led the GFSAD30SACE product generation effort. Ms. Zhong was instrumental in the design, coding, computing, analyzing, and synthesis of Landsat derived nominal 30-m GFSAD30SACE cropland product of the South American continent for the nominal year 2015. She was also instrumental in writing the manuscripts, ATBD, and user documentation.

Dr. Chandra Giri, Chief, Sensing and Spatial Analysis Branch, United States Environmental Protection Agency (USEPA), provided guidance and intellectual insights into South American cropland mapping. He is one of the co-I's of the project and was instrumental in developing framework of the GFSAD30SACE product, writing the manuscripts, ATBD, and user documentation.

Dr. Prasad S. Thenkabail, Research Geographer, United States Geological Survey, is the Principal Investigator (PI) of the GFSAD30 project. Dr. Thenkabail was instrumental in developing the conceptual framework of the GFSAD30 project and the GFSAD30SACE product. He provided guidance and intellectual insights throughout the GFSAD30 project and contributed in writing of the manuscripts, ATBD, and user documentation.

Dr. Pardhasaradhi Teluguntla, Research Scientist, Bay Area Environmental Research Institute (BAERI) at the United States Geological Survey (USGS), provided input and insights on cropland extent product generation for the South American continent. He made a significant contribution in writing the manuscripts, ATBD, user documentation. He also created the baseline cropland mask at 1-km which was instrumental in masking croplands *versus* non-croplands initially.

Dr. Russell G. Congalton, Professor of Remote Sensing and GIS at the University of New Hampshire, led the independent accuracy assessment of the entire GFSAD30 project including GFSAD30SACE.

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

Mr. Adam Oliphant, Geographer, United States Geological Survey, provided Google Earth Engine (GEE) cloud computing insights and support in generating the GFSAD30SACE cropland product of the South American continent.

Dr. Jun Xiong, Research Scientist, Bay Area Environmental Research Institute (BAERI) at the United States Geological Survey (USGS), provided Google Earth Engine (GEE) cloud computing insights and support in generating the GFSAD30SACE cropland product of the South American continent.

Ms. Varsha Vijay, PhD student @ Duke University, helped in algorithm development of the GFSAD30SACE cropland product of the South American continent. She also participated in several discussions on accuracy assessment.

Mr. Justin Poehnelt, former member of GFSAD30 project and Computer Scientist with the United States Geological Survey, contributed to the initial conceptualization and development of the <http://croplands.org> website.

II. Historical Context and Background Information

Monitoring global croplands is imperative for ensuring sustainable water and food security for the people of the world in the twenty-first century. However, the currently available cropland products suffer from major limitations such as: (1) the absence of precise spatial locations of the cropped areas; (2) coarse resolution of map products with significant uncertainties in areas, locations, and detail; (3) uncertainties in differentiating irrigated areas from rainfed areas; (4) absence of crop type information and cropping intensities; and/or (5) the absence of a dedicated Internet data portal for the dissemination of cropland products. Therefore, the Global Food Security-support Analysis Data (GFSAD) project aimed to address these limitations by producing cropland maps at 30m resolution covering the globe, referred to as Global Food Security Support-Analysis Data @ 30-m (GFSAD30) products. This Algorithm Theoretical Basis Document (ATBD) provides a basis upon which the GFSAD30 cropland extent product was developed for the continent of South America (GFSAD30SACE, Table 1).

South America is a traditional food producer and exporter, and it has the potential to continuously increase its food production in the near future given its relatively low population and rich availability of land and water resources. In the past 50 years, South America has experienced extensive cropland expansion (Graesser et al. 2015) and the fastest agricultural productivity growth in any developing region (Ludena 2010). South America has great potential to further expand and intensify crop production (Conforti 2011) due to availability of unexploited arable lands for cultivation (Fischer, van Velthuizen, and Nachtergaele 2011), water resources, labor forces, rich experiences in cultivation practices, and infrastructure (Zeigler and Truitt Nakata 2014). In the recent past, about 40 percent of agricultural production was contributed by cropland expansion for Gross Domestic Product (GDP) and according to FAO it will remain equally important in the future (Alexandratos and Bruinsma 2012). This is in sharp contrast to the rest of the world where the focus will be to increase the yield rather than cropland expansion.

However, expansion of cropland in South America does not come without serious negative consequences in providing ecosystem goods and services to environment and society. For example, South America has the largest area of tropical forest in the world with rich biodiversity (Myers et al. 2000). Cropland expansion occurs at the expense of wild lands such as forest, (Castiblanco, Etter, and Aide 2013; Gasparri, Grau, and Angonese 2013; Müller et al. 2012) grassland, and savanna (Baldi and Paruelo 2008; Graesser et al. 2015). This will have negative consequences in providing ecosystem goods and services to the entire world (Foley et al. 2007; Viglizzo and Frank 2006). For example, crop expansion accelerates carbon emission (Karstensen, Peters, and Andrew 2013), reduces sequestration due to loss of forests, and impacts biodiversity (Fearnside

2001). Thus, improving our scientific understanding of the cropland distribution and dynamics over space and time is critically important. The South American continent has experienced rapid changes in the last 50 years, which is likely to continue throughout the twenty-first century.

Despite this, accurate, reliable, and consistent information on the extent and distribution of croplands of the South American continent at finer spatial scale is unavailable. In order to fill this gap, the Global Food Security–Support Analysis Data at 30 m (GFSAD30) project implemented by the U.S. Geological Survey (USGS) together with a number of partner organizations (croplands.org) and funded by NASA aims to generate high resolution (nominal 30 meters) cropland products from multi-sensor remote sensing images (<http://geography.wr.usgs.gov/science/croplands/>). GFSAD30 has successfully developed a 1-km resolution cropland extent product and a watering sources product (Teluguntla et al. 2015) by synthesizing four cropland spatial distribution studies: Thenkabail et al. (2009), Pittman et al. (2010), Yu et al. (2013), Friedl et al. (2010). The research group is generating 30-meter spatial resolution cropland products (croplands.org) for the entire world. This paper is focused on development of cropland extent and areas of the South American continent at nominal 30-m spatial resolution using Landsat data.

Table 1. Basic information of the Global food security support-analysis data @ 30-m cropland extent product for the South American continent (GFSAD30SACE)

Product Name	Short Name	Spatial Resolution	Temporal Resolution
GFSAD 30-m Cropland Extent Product of South America	GFSAD30SACE	30-m	nominal 2015

III. Rationale for Development of the Algorithms

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

The two most common methods for land-cover mapping over large areas using remote-sensing images are manual classification based on visual interpretation and digital per-pixel classification. The former approach delivers products of high quality, such as the European CORINE Land Cover maps (Büttner, 2014). Although the human capacity for interpreting images is remarkable, visual interpretation is subjective (Lillesand et al., 2014), time-consuming, and expensive. Digital per-pixel classification has been applied for land-cover mapping since the advent of remote sensing and is still widely used in operational programs, such as the 2005 North American Land Cover Database at 250-m spatial resolution (Latifovic, 2010). Pixel-based classifications such as maximum likelihood classifier (MLC), neural network classification (NN), decision trees, Random Forests (RF), and Support Vector Machines are powerful, and fast classifiers that help differentiate distinct patterns of landscape.

Both supervised and unsupervised classification approaches are adopted in pixel-based classifiers. However, per-pixel classification includes several limitations. For example, the pixel's square shape is arbitrary in relation to patchy or continuous land features of interest, and there is significant spectral contamination among neighboring pixels. As a result, per-pixel classification often leads to noisy classification outputs – the well-known “salt-and-pepper” effect. There are other limitations of pixel-based methods: 1. they fail to fully capture the spatial information of high resolution imagery such as from Landsat 30-m imagery, and 2. they often, classify the same field (e.g., a corn field) into different classes as a result of within field variability. This may often result in a field with a single crop (e.g., corn) classified as different crops.

For the creation of the GFSAD30SACE data product, we applied the most commonly used supervised pixel-based classifier (Pelletier et al., 2016, Tian et al., 2016, Shi and Yang, 2015, Huang et al., 2010): Random Forests (RF's).

IV. Algorithm Description

We used the Random Forest classifier to classify croplands and non-croplands. Previous studies have shown that the RF classifier performs better in land cover classification (Gislason, Benediktsson, and Sveinsson 2006; Pal 2005; Rodriguez-Galiano et al. 2012) compared to the traditional maximum likelihood classifier (MLC). Furthermore, since Random Forest is a nonparametric classifier (Duda, Hart, and Stork 2012), it is appropriate and powerful for land cover classification in which remote sensing data collected from multiple sources are used (Gislason, Benediktsson, and Sveinsson 2006). As an ensemble classification model, multiple decision trees are trained in one Random Forest classification model, each of which uses a set of bootstrapped samples from the original samples. Each decision tree contributes a vote of class, and the final output class of the classifier is the majority vote of the trees. In each tree node, only a random subset of variables is selected for classification in order to reduce tree correlation (Breiman and Adele 2016).

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

a. Input data

i. Region Definition

The study area consists of the entire continent of South America, covering twelve sovereign countries (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Suriname, Uruguay, and Venezuela), and two territories including French Guiana and Falkland Islands. The continent's area is ~1.78 billion hectares with a total population of 418 million people. Forest covers half of the South America continent (Giri and Long 2014). The country boundaries were determined by the Global Administrative Unit layers (GAUL) of United Nations (<http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691&currTab=simple>).

ii. Reference Samples

Training samples for croplands *versus* non-croplands were generated through multiple steps. First, we created random points and drew homogenous polygons around the random point with the size of approximately three by three Landsat pixels (90-m by 90-m). We then defined the land cover category of the polygons as croplands, or non-croplands with visual interpretation assisted by (i) high-resolution images available from Google Earth, (ii) Landsat images, and (iii) a MODIS NDVI temporal profile. In the second stage, additional training samples were selected where classification performance was poor. The training sample selection and classification model training were performed tile by tile.

Reference training/testing data were obtained in the ways. First, we gathered random samples by interpreting sub-meter to 5-meter very high spatial resolution imagery (VHRI) throughout South America available to us from the National Geospatial Agency (NGA). There was a total of 3000+ samples from VHRI spread across South America. Second, some other global/region projects (Teluguntla et al., 2015, Thenkabail et al., 2012) shared their valuable reference datasets. To incorporate these reference data in our project, we converted their labeling system ("cross-walked") to be consistent with the labeling scheme of our project (Teluguntla, 2015).

iii. Satellite Imagery: Landsat data

We used Landsat 5-8 (Table 2) Top-of-atmosphere images provided by USGS acquired during January 1, 2013 to January 1, 2016. In most areas, we divided a calendar year into 6 periods starting with Julian day 1. The first 5 periods consisted of 60 days while the sixth period consisted of 65 days, for a total of 365 days. All images acquired in one period across 3 years were composited together to find the medium value of every pixel except in areas with high percentage of cloud coverage across the year, such as coastal Brazil and Ecuador. In these areas, only the last three periods (day 181 – day 365) – the dry periods – were used, or the year was divided into three periods with every 120 days being a period (125 for the last period), in order to acquire clear images for each period.

The following bands were selected from Landsat images to use in classification: blue, green, red, NIR, SWIR1, SWIR2, and NDVI (Table 2). We also add an NDVI standard deviation (NDVI SD) band which is calculated from the composited NDVI bands of all periods in a year. The

NDVI SD band reflects the fluctuation of NDVI within a year. Further, considering that topographic characteristics can often affect whether farming is possible, a slope band was added to the set of classification bands, and was calculated from the Shuttle Radar Topography Mission (SRTM) (Farr et al. 2007) digital elevation at 1 arc-sec (approximately 30m) resolution dataset. Overall, a total of 44 bands were used in the areas that were composited by 6 periods, and 23 bands in the areas that were composited by 3 periods.

Table 2. Characteristics of Landsat 5, 7 &8 data used in the study along with band indices (VI).

Band Name	Landsat 5 TM Spectral Range μm	Landsat 7 ETM+ Spectral Range μm	Landsat 8 OLI Spectral Range μm	VI Name	Equation
Blue	0.45-0.52	0.45-0.52	0.45 – 0.51	NDVI	$\frac{NIR - Red}{NIR + Red}$
Green	0.52-0.60	0.52-0.60	0.53 – 0.59		
Red	0.63-0.69	0.63-0.69	0.64 – 0.67	NDVI SD	$\sqrt{\frac{\sum(NDVI - \overline{NDVI})^2}{n - 1}}$
NIR	0.76-0.90	0.77-0.90	0.85 – 0.88		
SWIR1	1.55-1.75	1.55-1.75	1.57 – 1.65		
SWIR2	2.08-2.35	2.09-2.35	2.11 – 2.29		

Note: NIR = near infrared, SWIR = shortwave infrared, TM = thematic mapper, ETM+ = enhanced thematic mapper plus, OLI = Operational Land Imager, NDVI = normalized difference vegetation index, NDVI SD = standard deviation of NDVI.

b. Theoretical description

i. Definition of Croplands

For all products within GFSAD30, cropland extent was defined as, “lands cultivated with plants harvested for food, feed, and fiber, including both seasonal crops (e.g., wheat, rice, corn, soybeans, cotton) and continuous plantations (e.g., coffee, tea, rubber, cocoa, oil palms). Cropland fallows are lands uncultivated during a season or a year but are farmlands and are equipped for cultivation, including plantations (e.g., orchards, vineyards, coffee, tea, and rubber” (Teluguntla et al., 2015). Cropland extent includes all planted crops and fallow lands. Non-croplands include all other land cover classes other than croplands and cropland fallows (Figure 1).



Figure 1. Illustration of definition of cropland mapping. Croplands included: (a) standing crop, (b) cropland fallows, and (c) permanent plantation crops.

ii. Algorithm

The study used the pixel-based supervised classification machine-learning algorithm Random Forest (RF) to create the cropland extent product. The algorithm is described in detail below. South America was stratified into five separate refined FAO agro-ecological zones (Figure 2) to facilitate the optimal classification.

c. Practical description

i. Random Forest (RF) Algorithm

All supervised pixel-based classifications are heavily dependent on the input training samples selected. In order to discriminate croplands under various environments and conditions, the sample size of the initial training dataset needs to be large, especially in complex regions. All samples were selected to represent a 90-m x 90-m polygon. We used sub-meter to 5-m very high spatial resolution imagery to generate croplands *versus* non-cropland samples using multiple interpreters across South America. Approximately 3000+ data samples were used generated from these interpretations. These samples were used in training the Random Forest (RF) machine learning algorithm.

We processed and composited Landsat images, and performed Random Forest classification on the Google Earth Engine cloud computing platform. GEE stores all Landsat data seamlessly for the entire continent. Cloud computing offers the power of parallel processing across thousands of computers, thus allowing us to classify the entire continent in a matter of minutes.

Parameter settings of the Random Forest classifier are listed below:

- Number of decision trees to grow: 600. After testing on several training datasets, the number of 600 trees balances the classification error rate and computation load. Therefore, 600 trees were built in each Random Forest classifier.
- Variables per split: 2 The default setting of the Random Forest classifier algorithm, square root of the number of inputs, was accepted and therefore two randomly selected variables were used in each tree node to determine the class.

- Minimum leaf population: 1. The minimum size of a terminal leaf in the tree could be as small as one, meaning that no pruning was done.
- Fraction of input to bag per tree: 50% The Random Forest classifier uses an out-of-bag (oob) mode so that half of the original set of samples are randomly selected with replacement as input samples in the training of one decision tree, and the rest samples are used to estimate the oob error. More information about oob and oob error estimates is available at (Breiman and Adele 2016).

The RF algorithm was coded and run separately on each of the five distinct zones (Figure 2). The zones were delineated based on their agro-ecology, climate, elevation, and agricultural patterns. This was done based on the internal discussions within the team keeping in view the ease of applying the RF algorithm for best possible results in separating croplands from non-croplands. We followed an iterative process to run the Random Forest classifier for each of the 5 zones separately, until the results were satisfactory. For the land cover classes and areas that were poorly classified, we manually added training samples and re-ran the classifier.

After each round of classification, we identified areas where croplands were misclassified as non-croplands. In order to reduce omissions, we ran another round of classification on only the areas being classified as non-croplands. Afterwards, the croplands classified through multiple iterations were mosaicked together to generate an overall Cropland Extent Map for the continent. This process was adopted in each of the 5 zones (Figure 2).

To improve spatial coherency, the Cropland Extent Maps of each of the 5 zones were smoothed by performing majority filtering, twice. In this process, each pixel was reclassified into the majority land cover class of its eight surrounding pixels. The output from majority filtering is the final Cropland Extent Map (Figure 3).

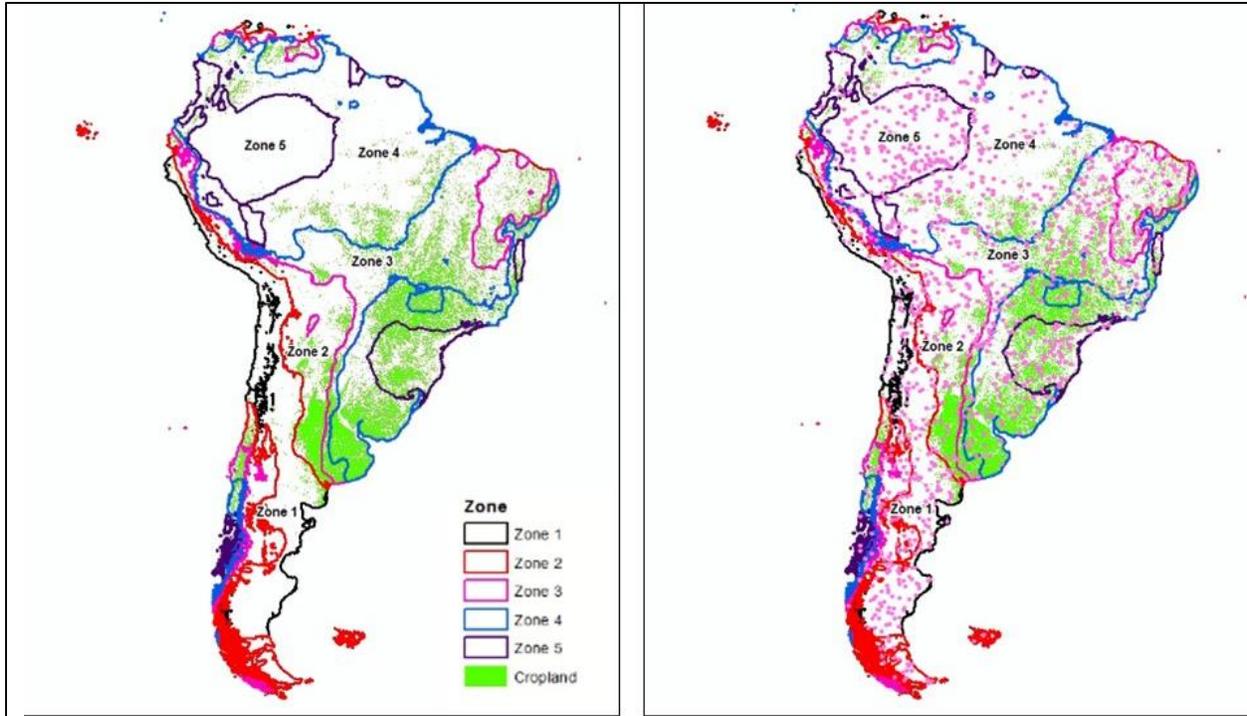


Figure 2. Study zones and reference training samples. The maps here show the delineation of 5 study zones. The RF algorithm was run separately in each of the 5 zones. The map on the right also shows the distribution of the reference training samples spread across the continent. These reference training samples were used in training the RF algorithm.

ii. Programming and codes

The pixel-based supervised machine learning algorithm (RF) was coded on Google Earth Engine (GEE) using Python and Java Scripts using an Application Programming Interface (API). The codes are available in a zip file and are available for download through LPDAAC along with this ATBD.

iii. Results

The study led to a cropland extent product @ nominal 30-m (Figure 3) for nominal year 2015 for the South American continent. This is referred to as global food security-support analysis data @ 30-m of South America: cropland extent product (GFSAD30SACE). Spatial distribution of these croplands can be visualized at: croplands.org. The full 30-m resolution of the product can be “zoomed-in” and visualized (e.g., Figure 4). A visual comparison of our Cropland Extent Map shows a similar distribution pattern of major croplands in South America as depicted by GlobCover 2009 and GCE V1.0 cropland maps (Figure 5). The total cropland area and cropland area in each country (Table 3) calculated from our Cropland Extent Map are also very close to the area calculated from GlobCover 2009 major cropland map and GCE V1.0 major cropland. Some regions where major differences of cropland distribution were observed are Chile, Colombia, Ecuador, and Venezuela. Our Cropland Extent Map generates a higher value of cropland area in Chile and lower value in Colombia compared with GlobCover 2009 major cropland map and

GCE V1.0 major cropland map, and much lower estimation of cropland area in Venezuela compared with GCE V1.0 major cropland map. Moreover, GlobCover 2009 major and minor cropland maps have a wider distribution of cropland than the three-cropland maps mentioned above. To achieve a more accurate evaluation and comparison of cropland areas based on cropland maps, sub-pixel cropland area should be calculated to replace full-pixel area across all products as proposed by Thenkabail et al. (2007). Sub-pixel areas provide actual areas, especially when computing areas from coarser resolution imagery such as GlobCover 2009. However, sub-pixel areas may not be needed to determine areas in 30-m product such as this study.

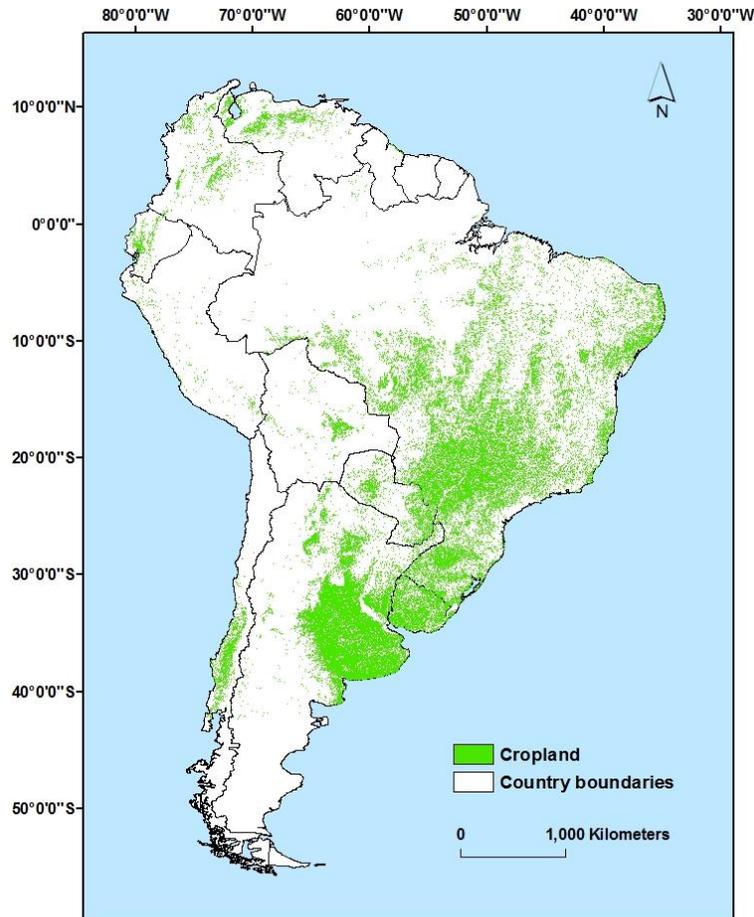


Figure 3. Cropland Extent Map of South America at a spatial resolution of 30-m for the years 2013 - 2015. This product is made available for visualization at: <https://croplands.org>.

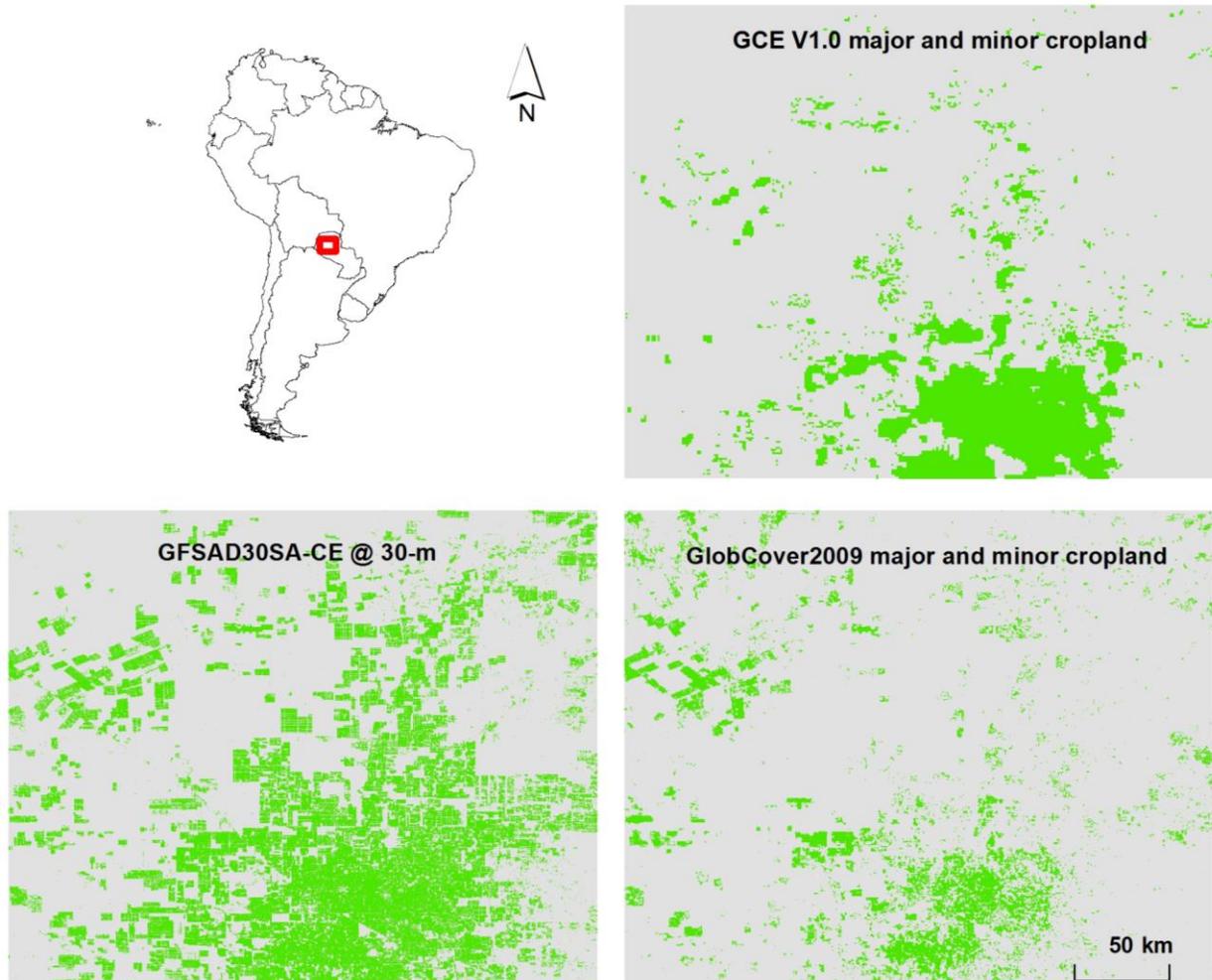


Figure 4. “Zoom-in” view of a selected area of the cropland Extent Map of South America at a spatial resolution of 30-m for the years 2013 - 2015. The South American product is made available for visualization at: [https:// croplands.org](https://croplands.org).

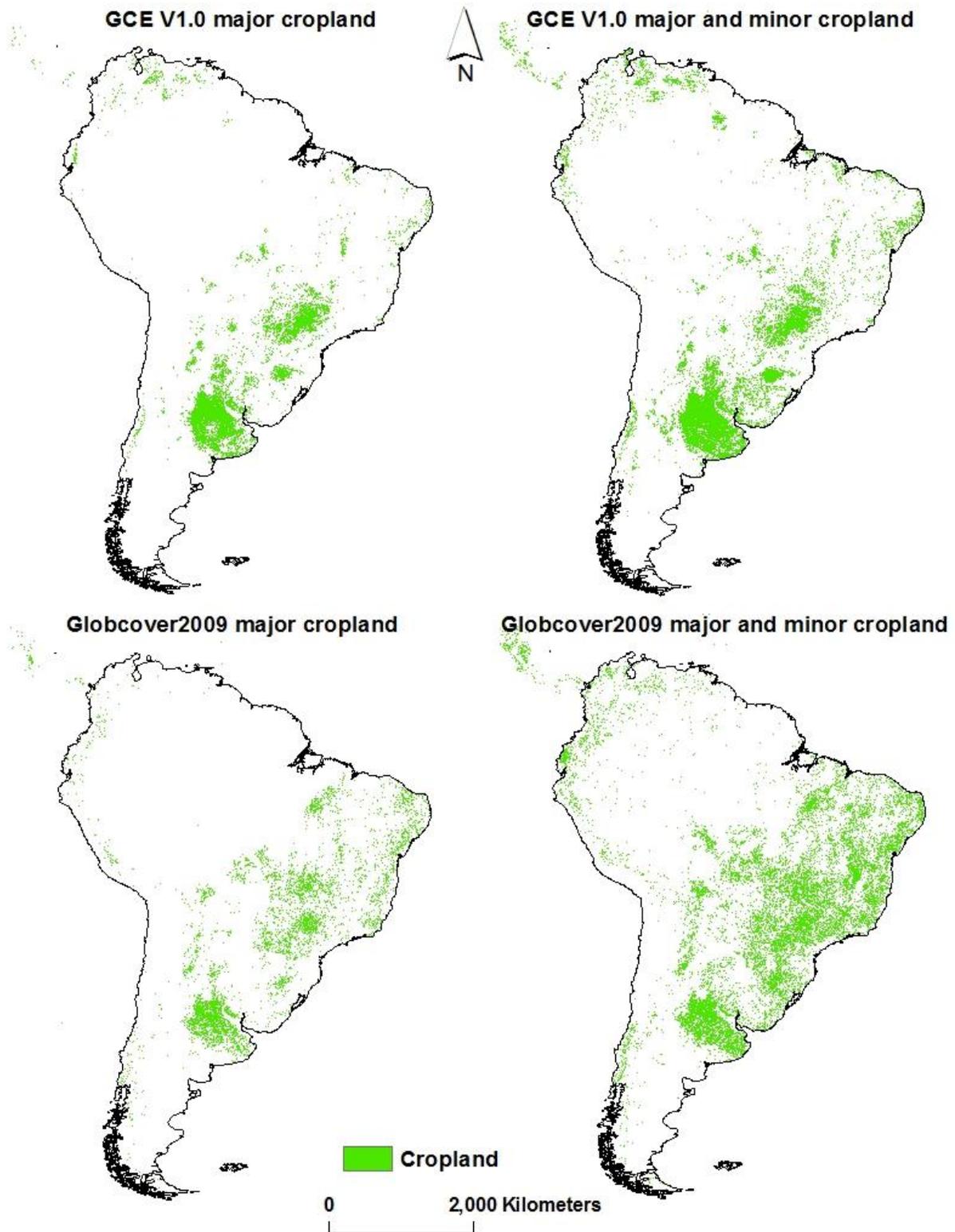


Figure 5. Comparison of GFSAD30SACE 30-m cropland product with three other coarser resolution cropland products.

iv. Cropland areas of South America

According to our Cropland Extent Map, the total net cropland area (TNCA) in South America is 151,994,479 hectares, or 8.7% of its total land area (Table 3). Brazil has the largest cropland area (42.1% of the total net cropland area of South America) followed by Argentina (25.3%). So, Brazil and Argentina have the largest cropland areas in South America relative to TNCA of the South American continent. However, as a % of the total geographic area of the country, Uruguay, and Paraguay have highest % areas of croplands with 66.9%, and 23.2% respectively. However, Uruguay, and Paraguay only have 7.7%, and 6.1% (Table 3) of total net cropland areas in South America.

The following major factors may contribute to prediction errors:

- (1) Limitations of the input Landsat data. Some areas are persistently covered by clouds, such as the Guayas region in Ecuador where large areas of plantation are present, resulting in low quality Landsat data (Giri and Long 2014) and therefore unsatisfactory classification results in these areas. Additionally, the anomaly of the Landsat 7 ETM + Scan Line Corrector caused problematic “stripes” in the Landsat 7 ETM + images. In future studies, more reliable remote sensing images such as Landsat 8 and Sentinel-2 should be used in the classification training and prediction.
- (2) Under-representation of landscape characteristics. Though training samples are selected to represent the diverse land cover types of South America comprehensively, some landscape may still be missed in the training dataset. Therefore, the input variable values of prediction cases may exceed the range of training cases, resulting in misclassification.
- (3) Uncertainties in the training dataset. We selected the training dataset and identified the cropland cover categories carefully. However, the visual interpretation of land cover may be incorrect in some cases, especially when reliable reference data in certain areas for the mapping period are not available. Therefore, errors are likely to exist in the training dataset which jeopardized the classification accuracy from the root. In order to increase accuracy of the training dataset, field survey data or high-resolution images and photographs should be used in training data generation.

Although our Cropland Extent Map agrees with other cropland maps in terms of the major spatial distribution pattern, discrepancies exist in some areas of the continent. Three major reasons may have contributed to the disagreements: 1. inconsistent definition of cropland class, 2. Maps produced at different time periods, and 3. different spatial resolution of the cropland maps. Inconsistency in cropland definition may contribute most to the disagreement between cropland maps. As mentioned above, pasture and rangeland are excluded from the ‘cropland’ class in our Cropland Extent Map, but included in the GlobCover 2009 agricultural classes. Therefore, the GlobCover 2009 major and minor cropland map in which Class 11, 14, and 20 of GlobCover 2009 were clumped into one ‘cropland’ class tends to over-estimate cropland areas as defined in this project.

In addition, the cropland maps were developed for different years, during which land cover changes may have happened; This would in turn result in disagreement of land cover descriptions even if the descriptions of cropland maps were all accurate for their mapping time. GlobCover

2009 was developed for the period of January – December 2009, and GCE V1.0 was synthesized from four cropland maps, the mapping time of which ranges widely from 2000 to 2010, while our Cropland Extent Map was developed for the years 2013 – 2015. From 2010 to 2013, some croplands might have been abandoned while new croplands were developed.

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

Country	Land Area ¹	GFSAD30 ²	MIRCA 2014 ³	FAO Agricultural land ⁴	GIAM-GMRCA ⁵	GRIPC 2005 ⁶	Percent of total Croplands	percent of total Land Area
Name	Ha	Ha	Ha	Ha	Ha	Ha	%	%
Brazil	845047923	63994709	58705445	68505500	91603674	102616446	42.10%	7.6%
Argentina	273879142	38383784	34778946	32034000	43623158	55168208	25.25%	14.0%
Uruguay	17502364	11709192	1567659	1910103	3735751	2879925	7.70%	66.9%
Paraguay	39733840	9238047	4868887	3908300	5567578	7481895	6.08%	23.2%
Venezuela	88065844	7630606	3766502	3402600	4151851	12681187	5.02%	8.7%
Colombia	111070496	6502398	4269633	3360660	5905473	17748936	4.28%	5.9%
Chile	74254717	5974811	3188325	1731620	3927135	3095281	3.93%	8.0%
Bolivia	108369501	3737749	3230784	3954078	9017920	3602161	2.46%	3.4%
Ecuador	24864686	2413259	3217439	2546492	3133011	4022652	1.59%	9.7%
Peru	127619048	2192154	5115162	4438080	5202729	2203578	1.44%	1.7%
Guyana	19705882	161870	563968	445550	280303	621627	0.11%	0.8%
Suriname	16200000	45464	106736	63666	108439	248805	0.03%	0.3%
Trinidad & Tobago	514286	10436	124446	46980	10449	150099	0.01%	2.0%
Total	1746827730	151994479	123503931	126347629	176267471	212520801	100%	8.7%

Note:

1= Total land area is land area excluding area under inland water bodies

2= GFSAD30 current study

3= Monthly irrigated and rainfed crop areas (MIRCA) around the year 2014 derived by Portman et al.

4= FAO Agricultural land area excluding pasture based on FAO2013 statistics consider nominal 2015

<http://www.fao.org/faostat/en/#data/QC>

5= Global croplands derived from Global Irrigated Area Mapping (GIAM) and

Global Map of Rainfed Cropland Areas (GMRCA) by Thenkabail et al., 2009 and Biradar et al., 2009

6= Global rain-fed, irrigated, and paddy croplands (GRIPC) derived by solmon et al., 2015

V. Calibration Needs/Validation Activities

In addition to spatial distribution, the Cropland Extent Map also indicates the cropland area in South America. We summarized the cropland area in each country and the total cropland area in South America from the Cropland Extent Map, and compared them with FAO statistics (FAO 2016) and the cropland area calculated from GlobCover 2009 major cropland map, GlobCover 2009 major and minor cropland map, GCE V1.0 major cropland map, GCE V1.0 major and minor cropland. It's noteworthy that the area here is full pixel area, meaning that the cropland pixels were considered as "pure" and the sub-pixel composition was not analyzed to extract fractional information of cropland (Thenkabail et al. 2007). Given that a Landsat pixel constitutes an area of 0.09 hectares, purity of the pixels can be trusted.

For this assessment, 1250 reference samples were used, that were collected independently of any reference training and testing samples used by the mapping team. Error matrices (Table 4) were generated for each of the five zones separately and for the entire South American continent providing producer's, user's, and overall accuracies (Story and Congalton, 1986, Congalton, 1991, and Congalton and Green, 2009).

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

Zone1		Reference Data			Total	User Accuracy
% TCASA=	1.8%	Crop	No-Crop			
Map Data	Crop	3	5	8	37.5%	
	No-Crop	3	239	242	98.8%	
Total		6	244	250		
Producer Accuracy		50.0%	98.0%		96.8%	

Zone2		Reference Data			Total	User Accuracy
% TCASA=	17.6%	Crop	No-Crop			
Map Data	Crop	25	8	33	75.8%	
	No-Crop	5	212	217	97.7%	
Total		30	220	250		
Producer Accuracy		83.3%	96.4%		94.8%	

Zone 3		Reference Data			Total	User Accuracy
% TCASA=	29.7%	Crop	No-Crop			
Map Data	Crop	39	9	48	81.3%	
	No-Crop	10	191	202	95.1%	
Total		49	200	250		
Producer Accuracy		79.6%	95.5%		92.4%	

Zone 4		Reference Data			Total	User Accuracy
% TCASA=	40.6%	Crop	No-Crop			
Map Data	Crop	40	11	51	78.4%	
	No-Crop	4	195	199	98.0%	
Total		44	206	250		
Producer Accuracy		90.9%	94.7%		94.0%	

Zone 5		Reference Data			Total	User Accuracy
% TCASA=	10.3%	Crop	No-Crop			
Map Data	Crop	21	6	27	77.8%	
	No-Crop	5	218	223	97.8%	
Total		26	224	250		
Producer Accuracy		80.8%	97.3%		95.6%	

All Zones		Reference Data			Total	User Accuracy
% TCASA=	100 %	Crop	No-Crop			
Map Data	Crop	128	39	167	76.7%	
	No-Crop	27	1,056	1,083	97.5%	
Total		155	1,095	1,250		
Producer Accuracy		82.6%	96.4%		94.7%	

Note: Total net cropland area of SA (TCASA) = 282.60Mha

Area weighted accuracy: 93.2%

For the entire continent, the weighted overall accuracy was 93.2% with producer's accuracy of 82.6% (errors of omissions of 17.4%) and user's accuracy of 76.7% (errors of commissions of 23.3%) (Table 5). When considering all 5 zones, the overall accuracies ranged between 92.4-96.8%, producer's accuracies ranged between 79.6-90.9%, and user's accuracies ranged between 76.7-81.3% (Table 4).

VI. Constraints and Limitations

GFSAD30SACE product mapped the croplands of South America @ nominal 30-m, which is the best-known resolution for cropland mapping over such a large area as the South American continent covering all countries. It also has high levels of accuracy with overall accuracy of 93.2%, producer's accuracy of 82.6%, and user's accuracy of 76.7%.

A producer's accuracy of 82.6% for the cropland class means an error of omission of 17.4%. This means 17.4% of the continental croplands were missing in the product. A user's accuracy of 76.7% for the cropland class for the continent means there is an error of commission of 23.3%. This means 23.38% of non-croplands are mapped as croplands. We tweaked the Random Forest algorithm (section IV) to maximize capturing as much cropland as feasible automatically. In this

process, some non-croplands were mapped as croplands as well. This is a preferred solution in order not to miss croplands or only to miss them minimally. As a compromise mapping some non-croplands as croplands becomes unavoidable.

There are numerous issues that cause uncertainties and limitations in the cropland extent product. Some of these issues are discussed here. First is temporal coverage. Five to ten-day Sentinel-2 and 16-day Landsat-8 coverage put together, there is substantial temporal coverage. Yet, achieving cloud-free or near cloud-free mosaics of the entire South American continent is difficult over some time periods (e.g., weekly, twice-monthly, monthly). This is not surprising given such a large area of the continent. If we were to have daily coverage over an area (e.g., like MODIS) then it becomes feasible to have more frequent (e.g., monthly or bi-monthly composites) temporal coverage of the continent that will help advance cropland mapping at improved accuracies. Second, there is a need for greater understanding of the Landsat-7 and Landsat-8 data on how well they are correlated in efforts to achieve better harmonization of data from the two different sensors. Third is the limitation of the reference training and validation data. In this project, we already have large training and validation data compared to any previous work as described in previous sections. Nevertheless, much wider and extensive field visits to different parts of the continent will be helpful in better understanding the issues involved, and as a result better mapping. For example, better understanding and defining of managed pastures from agricultural croplands are desirable. These and a better understanding of croplands through field visits as well as understanding of host of other issues (e.g., various types of irrigated and rainfed croplands, various types and ages of cropland fallows) will help improve cropland mapping. The greatest difficulties in cropland mapping in South America were in detecting, understanding, and delineating managed pastures from croplands. Furthermore, reduction in uncertainties in cropland mapping is feasible if we were to implement multiple machine learning algorithms rather than just Random Forests. These and numerous other issues will continue to persist in cropland mapping over such large areas as the South American continent. Nevertheless, advances made in this study are substantial, especially in developing a nominal 30-m cropland extent of the entire continent at very good accuracies.

VII. Publications

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

i. Peer-reviewed publications within GFSAD project

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

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

9781482217865 - CAT# K22125. Print ISBN: 978-1-4822-1786-5; eBook ISBN: 978-1-4822-1787-2. Pp. 678.

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

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

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

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

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

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

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

Web sites and Data portals:

<http://croplands.org> (30-m global croplands visualization tool)

<http://geography.wr.usgs.gov/science/croplands/index.html> (GFSAD30 web portal and dissemination)

<http://geography.wr.usgs.gov/science/croplands/products.html#LPDAAC> (dissemination on LP DAAC)

<http://geography.wr.usgs.gov/science/croplands/products.html> (global croplands on Google Earth Engine)
croplands.org (crowdsourcing global croplands data)

ii. Other relevant past publications prior to GFSAD project

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

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

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

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

Thenkabail, P.S., Biradar C.M., Noojipady, P., Cai, X.L., Dheeravath, V., Li, Y.J., Velpuri, M., Gumma, M., Pandey, S. 2007a. Sub-pixel irrigated area calculation methods. *Sensors Journal* (special issue: Remote Sensing of Natural Resources and the Environment (Remote Sensing Sensors Edited by Assefa M. Melesse). 7:2519-2538. <http://www.mdpi.org/sensors/papers/s7112519.pdf>.

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

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

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

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

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

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

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

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

iii. Books and Book Chapters

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

and Francis Inc. Press, Boca Raton, London, New York. ISBN 9781482217957 - CAT# K22130. Pp. 131-160

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

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

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

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

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

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

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

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

VIII. Acknowledgments

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

IX. Contact Information

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

Phone Number: 605-594-6116
Toll Free: 866-573-3222 (866-LPE-DAAC)
Fax: 605-594-6963
Email: lpdaac@usgs.gov
Web: <https://lpdaac.usgs.gov>

For the Principal Investigators, feel free to write to:

Prasad S. Thenkabail at pthenkabail@usgs.gov

For 30-m cropland extent product of South America, please contact:

Pardhasaradhi Teluguntla at pteluguntla@usgs.gov

Prasad S. Thenkabail at pthenkabail@usgs.gov

Ying Zhong at ying.zhong105@gmail.com

Chandra Giri at Giri.Chandra@epa.gov

More details about the GFSAD project and products can be found at: globalcroplands.org

X. Citations

Zhong, Y., Giri, C., Thenkabail, P.S., Teluguntla, P., Congalton, R.G., Yadav, K., Oliphant, A.J., Xiong, J., Poehnelt, J., Smith, C. (2017). *NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) Cropland Extent 2015 South America 30 m V001* [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MEaSUREs/GFSAD/GFSAD30SACE.001

XI. References

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

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

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

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

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

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

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

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

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

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

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

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

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

Hollstein, A., Segl, K., Guanter, L., Brell, M., Enesco, M., 2016. Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. *Remote Sensing* 8, 666.

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

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

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

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

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

- Lambert, M.-J., Waldner, F., Defourny, P., 2016. Cropland Mapping over Sahelian and Sudanian Agrosystems: A Knowledge-Based Approach Using PROBA-V Time Series at 100-m. *Remote Sensing* 8, 232.
- Languille, F., Déchoz, C., Gaudel, A., Greslou, D., Lussy, F. de, Trémas, T., Poulain, V., 2015. Sentinel-2 geometric image quality commissioning: first results, in: Bruzzone, L. (Ed.), *SPIE Remote Sensing*. SPIE, pp. 964306–964306–13.
- Latifovic, R., Homer, C., Ressler, R., Pouliot, D., Hossain, S.N., Colditz, R.R., Olthof, I., Giri, C. and Victoria, A., 2010. North American land Change Monitoring System (NALCMS). *Remote sensing of land use and land cover: principles and applications*. CRC Press, Boca Raton.
- Lillesand, T., Kiefer, R.W. and Chipman, J., 2014. *Remote sensing and image interpretation*. John Wiley & Sons.
- Löw, F., Schorcht, G., Michel, U., Dech, S., Conrad, C., 2012. Per-field crop classification in irrigated agricultural regions in middle Asia using Random Forest and support vector machine ensemble, in: Michel, U., Civco, D.L., Ehlers, M., Schulz, K., Nikolakopoulos, K.G., Habib, S., Messinger, D., Maltese, A. (Eds.), *SPIE Remote Sensing*. SPIE, pp. 85380R–85380R–11.
- Malinverni, E.S., Tassetti, A.N., Mancini, A., Zingaretti, P., Frontoni, E., Bernardini, A., 2011. Hybrid object-based approach for land use/land cover mapping using high spatial resolution imagery. *International Journal of Geographical Information Science* 25, 1025–1043.
- Motha, R.P., Leduc, S.K., Steyaert, L.T., Sakamoto, C.M., Strommen, N.D., Motha, R.P., Leduc, S.K., Steyaert, L.T., Sakamoto, C.M., Strommen, N.D., 1980. Precipitation Patterns in West South America. [http://dx.doi.org/10.1175/1520-0493\(1980\)108<1567:PPIWA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1980)108<1567:PPIWA>2.0.CO;2) 108, 1567–1578.
- Myint, S.W., Myint, S.W., Gober, P., Brazel, A., Gober, P., Brazel, A., Grossman-Clarke, S., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment* 115, 1145–1161.
- Pal, M., 2007. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. *International Journal of Remote Sensing* 27, 2877–2894.
- Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment* 187, 156–168.
- Rembold, F., Carnicelli, S., Nori, M., Ferrari, G.A., 2000. Use of aerial photographs, Landsat TM imagery and multidisciplinary field survey for land-cover change analysis in the lakes region (Ethiopia). *International Journal of Applied Earth Observation and Geoinformation* 2, 181–189.

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

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

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

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

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

Story, M. and R. Congalton. 1986. Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing*. Vol. 52, No. 3. pp. 397-399.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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