DCN
Version 1.0

September 2017
Version

NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) @ 30-m for Australia, New Zealand, China, and Mongolia: Cropland Extent Product (GFSAD30AUNZCNMOCE)

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	September 2017	Modification made according to USGS reviewer comments
1.2	September 2017	Modifications made according to LPDACC reviewer comments

CONTENTS

I. Members of the team

This Global Food Security-support Analysis Data 30-m (GFSAD30) Cropland Extent-Product of Australia, New Zealand, China, and Mongolia (GFSAD30AUNZCNMOCE) was produced by the following team members. Their specific role is mentioned.

Dr. Pardhasaradhi Teluguntla, Research Scientist, Bay Area Environmental Research Institute (BAERI) at United States Geological Survey (USGS) led the GFSAD30AUNZCNMOCE product generation effort. Dr. Teluguntla was instrumental in the designing, coding, computing, analyzing, and synthesis of the Landsat-8 derived nominal 30-m GFSAD30AUNZCNMOCE cropland product of Australia, New Zealand and China for the nominal year 2015. He was also instrumental in writing the manuscripts, ATBDs, and user documentations.

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 GFSAD30AUNZCNMOCE product. He made significant contribution in writing the manuscripts, ATBDs, User documentations, and providing scientific guidance throughout the GFSAD30 project.

Dr. Jun Xiong, Research Scientist, Bay Area Environmental Research Institute (BAERI) at United States Geological Survey (USGS), participated in the intellectual discussions and in provided inputs and insights on GFSAD30AUNZCNMOCE 30-m cropland extent product generation and shared his expertise in cloud computing.

Dr. Murali Krishna Gumma, Senior Scientist at the International Crops Research Institute for the Semi-Arid Tropics, helped collect reference data used in the machine learning algorithms.

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 GFSAD30AUNZCNMOCE 30-m cropland extent product of Australia, New Zealand and China.

Mr. Adam Oliphant, Geographer**,** United States Geological Survey (USGS), shared his expertise in cloud computing and Random Forest algorithm implementation in Google Earth Engine (GEE) for GFSAD30AUNZCNMOCE 30-m cropland extent product generation.

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

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

Dr. Temuulen T. Sankey, Assistant Professor, Northern Arizona University, produced the Mongolia cropland product and also performed its validation.

Ms. Aparna Phalke, PhD student, University of Wisconsin, shared cloud-computing expertise.

Ms. Corryn Smith, Student developer, helped in development of the croplands.org website.

II. Historical Context and Background Information

Monitoring global croplands (GCs) is imperative for ensuring sustainable water and food security for the people of the world in the Twenty-first Century. However, currently available cropland products suffer from major limitations such as: (a) Absence of precise spatial location of the cropped areas; (b) Coarse resolution nature of the map products with significant uncertainties in areas, locations, and detail; (c) Uncertainties in differentiating irrigated areas from rainfed areas; (d) Absence of crop types and cropping intensities; and (e) Absence of a dedicated web\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) product.

This Algorithm Theoretical Basis Document (ATBD) provides a basis upon which the GFSAD30 cropland extent product was developed for the countries of Australia, New Zealand, China, Mongolia (GFSAD30AUNZCNMOCE, Table 1), produced using Landsat-8 and Landsat-7 time-series satellite sensor data. This document provides comprehensive details of the GFSAD30AUNZCNMOCE production scheme that includes remote sensing data, reference and validation data, approaches, methods, machine learning algorithms, product generation, accuracy assessments, and area calculations.

It must be noted that cropland mapping for Mongolia was conducted separately by Dr. Sankey. A separate ATBD and user guide for Mongolia was not necessary as Mongolia only has about 0.06% to the global cropland areas, the methods and approaches mostly follow the one's described in this manuscript as well as in other continental studies.

Table 1. GFSAD30CE Products basic information for Australia, New Zealand, and China.

Product Name	**Short Name**	**Spatial resolution**	**Temporal coverage**
GFSAD30-m cropland Extent Product of Australia, New Zealand, China, and Mongolia	GFSAD30AUNZCNMOCE	30-m	Nominal 2015

Note: Nominal here means that the Landsat-8 16 day data used to produce the product is for two to three years (2013-2015), but the product is reported as nominal year 2015.

III. Rationale for Development of the Algorithms

Mapping the precise location of croplands enables the extent and area of agricultural lands to be more effectively captured, which is of great importance for managing food production systems and to study their inter-relationships with water, geo-political, socio-economic, health, environmental, and ecological issues (Thenkabail et al., 2010). Further, accurate development of all higher-level cropland products such as crop watering method (irrigated or rainfed), cropping intensities (e.g., single, double, or continuous cropping), crop type mapping, cropland fallows, 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 maps affect the quality of all higher-level cropland products reliant on an accurate cropland extent base map. However, precise and accurate cropland extent maps are currently nonexistent at the continental extent at a high spatial resolution (30-m or better). This lack of crop extent maps is particularly true for complex, small-holder dominant agricultural systems. By mapping croplands at a high spatial resolution at the continental scale, the GFSAD30project 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 classification methods: 1. they fail to 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.

We used supervised pixel-based classifier Random Forest (RF), which has been widely used in agricultural cropland studies over the years (Myint et al., 2011) and which is considered powerful and an ideal machine learning algorithm (Tian et al., 2016, Shi and Yang, 2015, Huang et al., 2010). A description of how to classify cropland extent of Australia, New Zealand, and China is provided in section 2.3 and its sub-sections.

This document describes, in detail, the development of the Global Food Security-support Analysis Data (GFSAD) @ 30-m for Australia, New Zealand, China, and Mongolia: Cropland Extent Product (GFSAD30AUNZCNMOCE). The approach involves the use of a supervised Random Forest (RF) algorithm to retrieve crop extent results from pixel-based classification (see overview of the methodology in Figure 1).

IV. Algorithm Description

An overview of the algorithm description is provided in Figure 1. The methodology used in this project (Figure 1) is briefly described in this paragraph to provide an overview of methods presented in detail in sub-sequent sections of this ATBD document. The process (Figure 1) involved combining year 2014-2015 16-day time-series Landsat-8 30-m data along with Landsat-7 30-m data. The process included several well designed steps (Figure 1). First, the data were pre-processed by cloud mask and gap filling on Google Earth Engine (GEE). Second, median value composites were created for 4 to 6 periods based on cloud-free or near-cloud-free wall-to-all coverage. Such a seasonal mosaic aided in achieving cloud free clear images of the continent. Each composite mosaic contained 8 bands as listed in Figure 1. Third, reference data were generated throughout the study area to train the RF algorithms. There are 3348 reference samples for this purpose. Fourth, the result of the pixel-based RF algorithm was to obtain the cropland extent product for Australia, New Zealand, and China. Fifth, the cropland product of Australia, New Zealand, and China was evaluated for accuracy using 3372 test samples. The process was iterated until adequate accuracies were attained. Accuracy assessments were performed by Dr. Russell Congalton and his PhD student, Kamini Yadav, independent of the production team. In this process, the validation data were only available to the accuracy assessment team and were hidden from the production team. As a result, there was completely independent accuracy assessment. Finally, the GFSAD30AUNZCNMOCE product was made available on croplands.org.

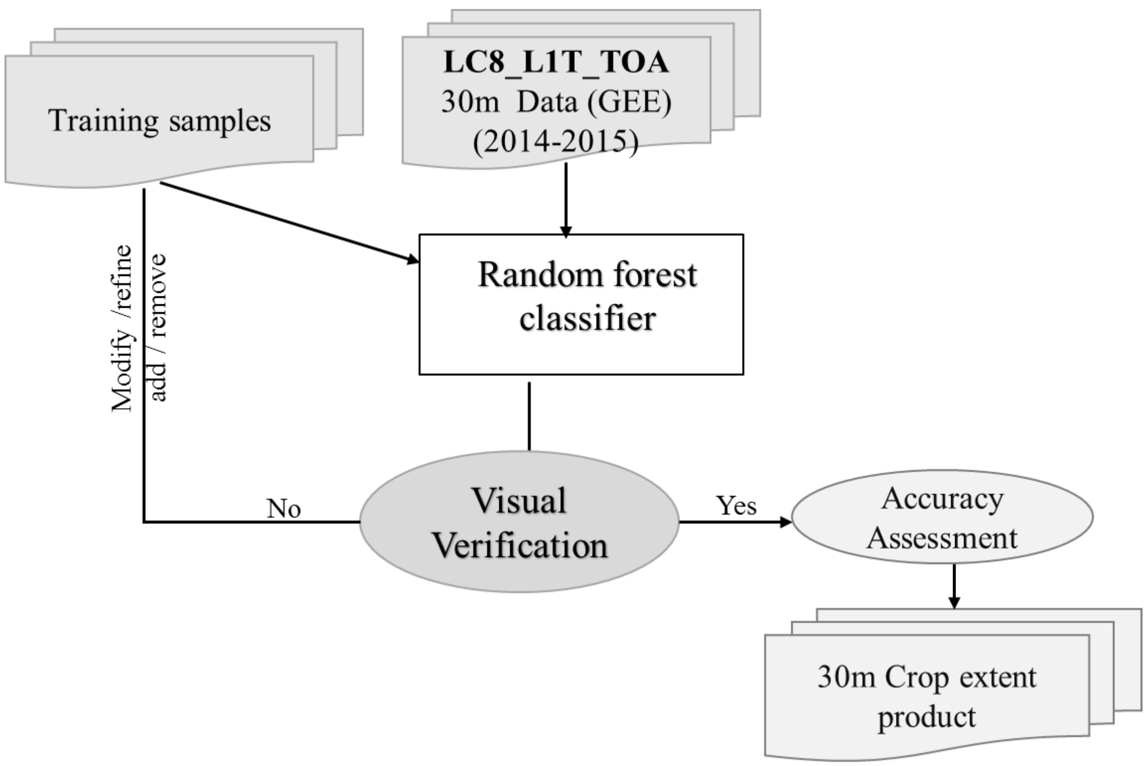

Figure 1. Flowchart of mapping methods for Landsat-8 derived cropland extent-product of Australia, New Zealand, and China for the nominal year 2015.

a. Input data

i. Region Definition

The study was conducted for the countries Australia, New Zealand and China (see Figure 2, Figure 5, Figure 6, and Figure 7). 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 Croplands Samples

Reference data are required for both training the machine learning algorithms (see section 2.3) as well as for validating the final products. First, we conducted an extensive field survey during September and October of 2014, the peak crop-growing season for crops in Australia. More than 4000 ground samples were collected from New South Wales (NSW), Victoria (VIC), South Australia (SA), and Western Australia (WA) regions of Australia following the specific guidelines on collecting ground reference data (Congalton, 2015). The sampling sites included various crop fields: such as Cereal crops (Wheat, Barley, and Oats), Legumes (Lupin, Lentils, Peas, and Beans), Oilseeds (Canola), Vegetables, Continuous crops (Orchard crops), Fodder crops (Alfalfa and sown pastures) and some fallow lands. Similarly, we obtained ground reference data for China through collaboration with the Chinese Academy of Agricultural Sciences (CAAS) who collecting field data that were spatially well spread out throughout agricultural cropland areas of China. The field survey gathered a total of 2120 ground samples including: (1) Location of samples (GPS position, location name, date of collection); crop properties (2) Croplands versus non-croplands; (3) irrigated or rainfed; (4) Crop intensity (single, double, triple, continuous cropping in 12 months); (5) Crop type (major crop types mentioned above, others); and (6) Digital photographs of each sample.

The ground data samples were collected from three main sources.

First, field surveys (or ground data) were collected during 2014 and 2015. The field-surveyed data were divided into three independent datasets with each set containing 1/3rd of the total samples (e.g., Table 2). The first set was used for training machine-learning algorithms (e.g., Random Forest). The second set was used to test the product. The third set was set aside and was used for independent accuracy assessment. In addition, we obtained reference-training data from the following reliable sources in addition to our own field data collections.

Second, random samples were obtained by interpreting sub-meter to 5-meter very high spatial resolution imagery (VHRI) data throughout Australia, New Zealand, and China available to US Government entities through the sub-meter to 5-m imagery obtained from the National Geospatial Agency (NGA). For this, we collected 1420 reference and 2710 validation samples.

Third, reference data were obtained from other reliable sources such as Geo Science Australia, Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES). The reference training data were used to "train" the Random Forest algorithm to separate croplands from non-croplands. This required us to keep adding training samples until optimal classification results were obtained (see section 2.3 and its sub-sections). A total of 2130 representative samples were used to "train" and separate croplands from non-croplands in China, 958 samples in Australia, and 260 samples in New Zealand (see Figure 2 showing the distribution of these samples) (Table 3).

The whole set of reference data including primary and secondary data were made available, at the following website: **https://croplands.org/app/data/search.**

Table 2. Ground data samples over Australia for the year 2014. The ground data samples were randomly split into reference or training (N1=1458 samples), validation set 1 for testing accuracies (N2=1488), independent validation set 2 for testing accuracies (N3=1465).

Code	Crop Description	Reference or training N1	Validation Set#1 N2	Validation Set#2 N3	Total number N
1	Alfafa	4	3	4	**11**
2	Barley	154	153	154	**461**
3	Beans	30	29	29	**88**
4	Canola	186	185	186	**557**
5	Lentils	65	65	65	**195**
6	Lupin	34	33	35	**102**
7	Oats	73	72	73	**218**
8	Peas	27	26	26	**79**
9	Wheat	283	283	284	**850**
10	Orchards	55	60	54	**169**
11	Sown-pasture	95	98	96	**289**
12	Season-2 Crops	20	16	18	**54**
13	Crop-harvested	9	9	10	**28**
14	Vegetables	1	1	1	**3**
15	Plantation	4	3	3	**10**
16	Cropland,others	30	29	25	**84**
20	Grazing/pastures	118	117	117	**352**
30	Non croplands	145	165	155	**465**
40	Fallow	125	141	130	**396**
	Total	**1458**	**1488**	**1465**	**4411**

Table 3. Number of reference samples used for training the Random Forest (RF) machine-learning algorithm and number of validation samples used for independent accuracy assessment.

Country	Class	Training samples	Validation samples
Australia	Crop	530	80
	No Crop	428	820
	Total	**958**	**900**
NewZealand	Crop	114	120
	No Crop	146	380
	Total	**260**	**500**
China	Crop	1346	340
	No Crop	784	1632
	Total	**2130**	**1972**

Note: The number of training and validation samples depended on the results. When optimal results obtained, we stopped adding further samples. The process requires starting with a certain sample number initially and progressively increasing sample number until optimal accuracies are reached.

iii. Image Stratification

The cropland versus non-cropland classification was carried out using the Random Forest (RF) machine-learning algorithm by stratifying the study area into refined agro-ecological zones (AEZs) (Figure 2). The AEZs were developed by the United Nation's Food and Agricultural Organization (UN FAO). However, this results in too many zones (which is not necessary given many zones have only a very small proportion of crops). Therefore, we combined some of these zones into broader refined AEZs (RAEZs) based on the convenience, and speed of applying the RF algorithm. This resulted in six broad RAEZs across China, Australia, and New Zealand (Figure 2). RF algorithm were trained for separating croplands from non-croplands in each of these RAEZs (Figure 2) using the reference training data falling within these zones. Working within each RAEZ also helped in data management and classification speed.

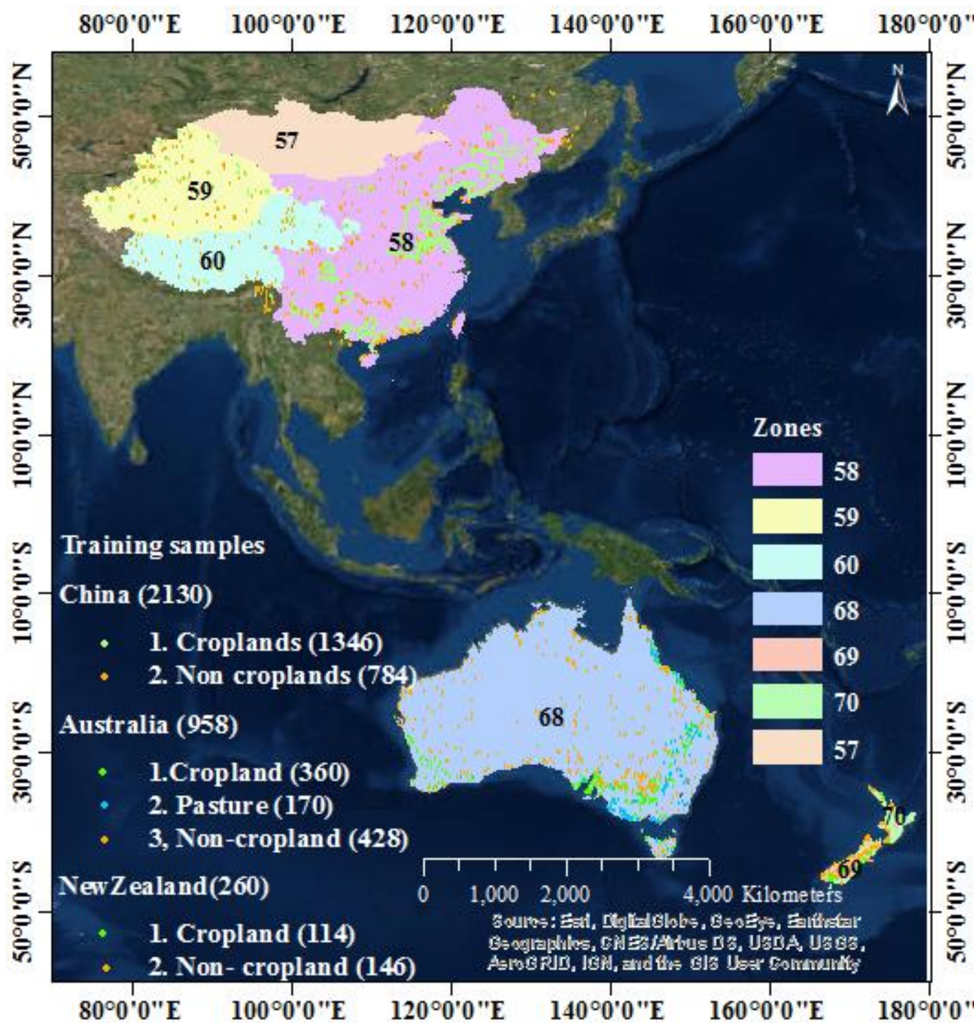

Figure 2. Stratification of the study area into distinct and broad refined agro-ecological zones (RAEZs). The figure also shows the distribution of the reference training data used in the Random Forest (RF) machine-learning algorithm. The Random Forest (RF) pixel-based supervised machine learning algorithm used in this study was "trained" using reference training data falling within each of these zones to separate croplands from non-croplands.

iv. Satellite Imagery: Landsat-8 and Landsat-7

In order to cover crop dynamics in different periods, Landsat-8 OLI (Roy et al., 2014) satellite data have been used for Australia, New Zealand, and China. In addition, Landsat-7 has been used to fill some data gaps (Irons et al., 2012) for China, aiming to provide seamless 30-m data for all time periods. There is a 16-day revisit time per Landsat-8 OLI and Landsat-7 ETM+ 30-m data. It is difficult to get continuous 8 to 16-day cloud free time-series data for wall-to-wall coverage for any part of the region. To overcome this limitation and to ensure cloud-free or near-cloud-free wall-to-all coverage, bi/tri-monthly composites, depending on the cloudiness of the countries\regions, were composed (e.g., Figures 3a and 3b). Finally, 30-m mega-file data-cubes (MFDCs) were created as per the following steps leading to a 48-band MFDC (Figure 3a) for Australia and New Zealand from 6 periods and a 32-band MFDC for China (Figure 3b) from 4 periods. A systematic detail of the MFDC composition is described below.

The goal of the time-composites was to achieve cloud-free or near cloud-free wall-to-wall composites over the entire study area (e.g., Figures 3a and 3b). This we wanted to achieve, using as many time-periods as possible as to get temporal stacks that can monitor phenology. However, the time-periods are decided by the ability to achieve cloud-free or near cloud-free images over a time-period. In Australia and New Zealand (Figure 3a) we were able to achieve the cloud-free or near cloud-free images at much shorter time-periods leading to 6 time-periods (period1: Julian days 1-60, period 2: Julian days 61-120, period 3: Julian days 121-180, period 4: Julian days 181-240, period 5: Julian days 241-300, period 6: Julian days 301-365; Figure 3a). In comparison, China (Figure 3b) required longer time-periods to achieve cloud-free or near cloud-free wall-to-wall coverage due to greater number of cloudy days over the country. As a result, there were 4 periods (period1: Julian days 1-90, period 2: Julian days 91-180, period 3: Julian days181-270, period 4: Julian days 271-365; Figure 3b).

The process involved gathering all the Landsat-8 16-day images over Australia and New Zealand (Figure 3a), and all the Landsat-8 as well as Landsat-7 images over China (Figure 3b) available for each time-period (e.g., period 1: 1-60 Julian days for Australia and New Zealand), and compositing each of the 8 bands by taking median value of each pixel of each band. These composites are called median value composites for each period for each band. The eight bands used in this study were (Figure 3a, 3b): blue (0.45-0.51μm), green (0.53-0.59μm), red 0.63-0.69μm), NIR (0.85-0.89μm), SWIR1 (1.55 1.65μm), SWIR2 (2.1-2.3 μm), and TIR1 (10.60-11.19μm) bands along with Normalized Difference in Vegetation Index (NDVI). Thereby, for Australia and New Zealand, eight median value bands composed over 6 time-periods resulted in a 48 band MFDC (Figure 3a). Whereas for China, eight median value bands composed over 4 time-periods resulted in a 32 band MFDC (Figure 3b). The band stack, and time-periods leading to MFDC are shown in Table 4 as well as in Figures 3a and 3b. All compositions were performed on the Google Earth Engine (GEE) cloud-based geospatial platform for planetary-scale data analysis (Gorelick et al,, 2017). Landsat top of atmosphere (TOA) products were used instead of surface reflectance (SR) due to the limited temporal availability of Landsat-7 and Landsat-8 surface reflectance imagery on GEE.

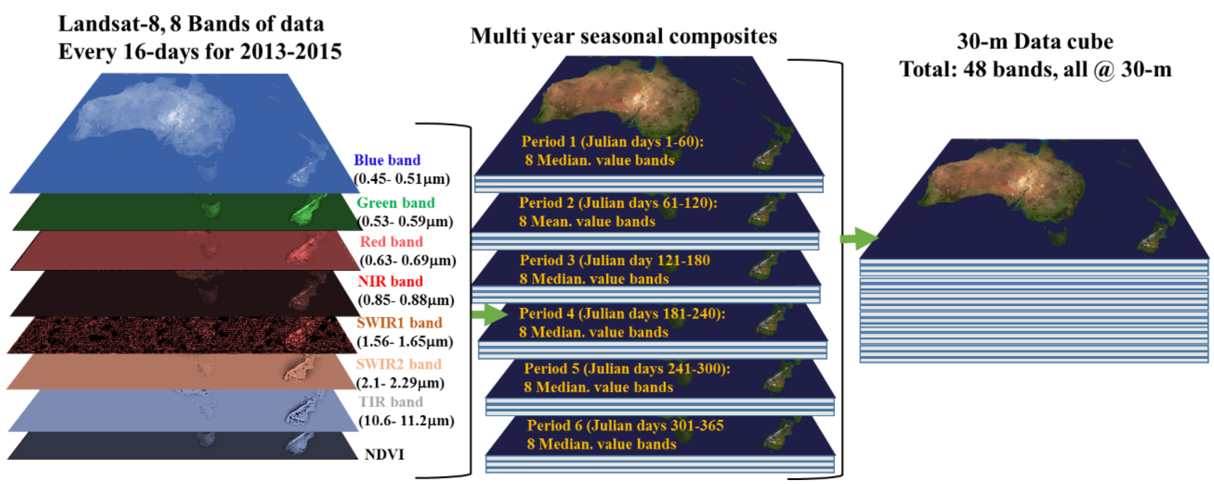

Figure 3a. 30-m Data-cube for the Australia and New Zealand regions composited for 6 time-periods using 2013-2015 Landsat-8 data. For each period (e.g., period 1: Julian days 1-1-60), eight bands (blue, green, red, NIR, SWIR1, SWIR2, TIR1, and NDVI Landsat-8) were composited, taking median value of a given pixel over the period 1. From 6 periods, there was a 48 band mega file data cube (MFDC).

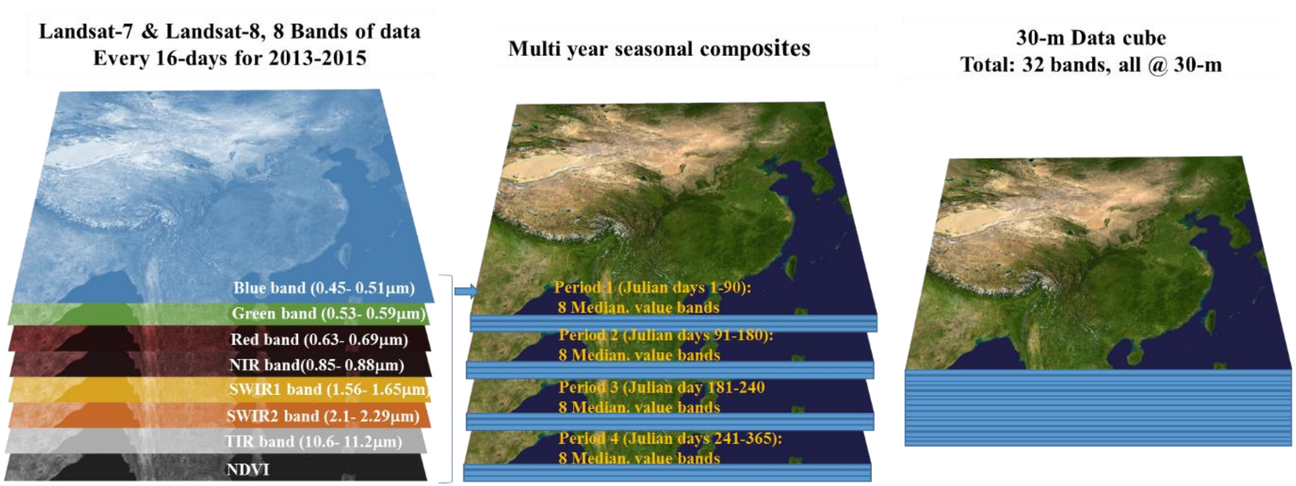

Figure 3b. 30-m Data-cube for China composited for 4 time-periods (e.g., period 1: Julian days 1 to 90) using every 16-day data of Landsat-8 and Landsat-7 for the years 2013-2015. For each period eight bands (blue, green, red, NIR, SWIR1, SWIR2, TIR1 and NDVI Landsat-8) were composited, taking median value of a given pixel over the period. From the 6 periods, there was a 32 band MFDC.

Table 4. The process of mega file data cube (MFDC) composition for the study areas based on median value composition of 8 Landsat-8 and\or Landsat-7 bands over 2013-2015 for 4 to 6 time-periods.

Region/ Country	Landsat products/ image Series	Years of data	Time-com-posited*	Bands used**	Mega-file Data Cube	Data Provider
name	**satellite, sensor**	**Years**	**Julian days over which Landsat data are time-compo-sited**	**# of bands for each**	**Total # of bands in MFDC**	**name**
Australia & New Zea-land	Landsat-8	2014 & 2015	C1:1-60 C2:61-120 C3:121-180 C4:181-240 C5:241-300 C6:301-365	Blue, green, Red, NIR , SWIR-1, SWIR-2, TIR-1 and NDVI	48	USGS
China & Mongolia	Landsat-7 & Landsat-8	2013, 2014 & 2015	C1:1-90 C2:91-180 C3:181-270 C4:271-365	Blue, green, Red, NIR , SWIR-1, SWIR-2, TIR-1 and NDVI	32	USGS

* C1:1-60 = composite 1 over Julian dates 1 to 60. Given Landsat-8 is acquired over every 16 days, there will be 4 images in first 60 days.

Then each band (e.g., blue) is derived using maximum value from these 4 images. Similarly for all bands.

similarly composite 2 C2:61-120 = taking images available during Julian day 61 and 120

**NIR - near-infrared, SWIR = short-wave infrared, TIR-1= thermal infrared

NDVI = normalized difference vegetation index,

b. Theoretical Description

i. Definition of Croplands

For all products within GFSAD30 cropland extent map, 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, and oil palms). Cropland fallows are lands uncultivated during a season or a year but are farm-lands and are equipped for cultivation, including plantations (e.g., orchards, vineyards, coffee, tea, and rubber)" (Teluguntla et al., 2015). Cropland extent also includes areas equipped for crop-ping but may not be cropped in a particular season or year. These are cropland fallows. So cropland extent includes all planted crops plus cropland fallows. Non-croplands include all other land cover classes other than croplands and cropland fallows.

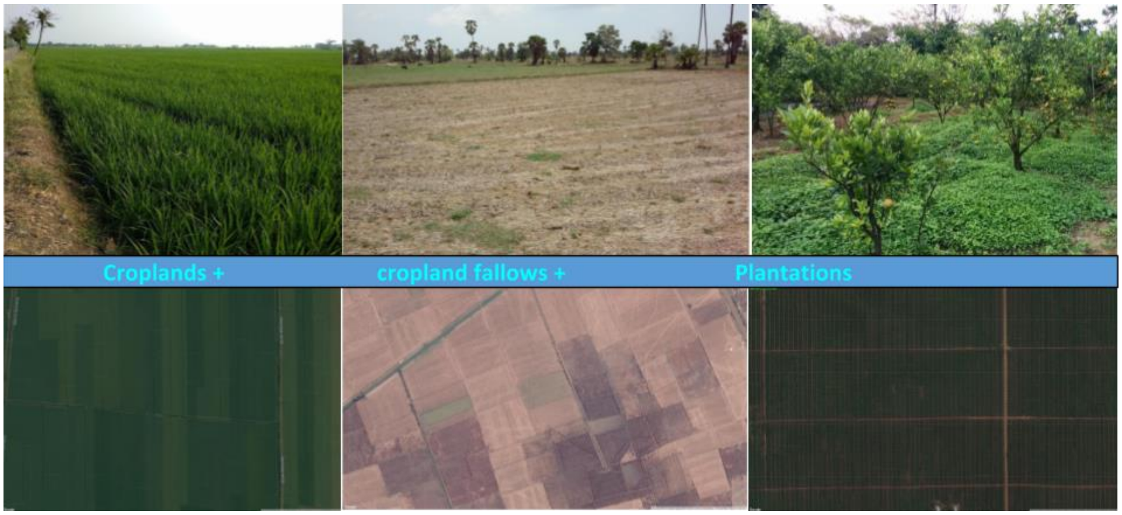

Figure 4. Illustration of definition of cropland mapping. Croplands included: (a) standing crop, (b) cropland fallows, and (c) permanent plantation crops. Note: + sign means adding. Means total net croplands= standing crops + cropland fallows + plantations.

ii. Algorithm

The study used one machine-learning algorithm to create the cropland extent product, which is the pixel-based supervised classifier Random Forest (RF). The algorithm is described in detail below. New Zealand was stratified into two separate refined FAO agro-ecological zones, China stratified into three zones and Australia as a single zone (Figure 2) to facilitate the optimal classification.

c. Practical Description

i. Random Forest Classifier (RF)

The Random Forest classifier is more robust, relatively faster, and easier to implement than many other classifiers (Pelletier et al., 2016). The Random Forest classifier uses bootstrap aggregating (bagging) to form an ensemble of decision trees by searching random subspaces from the given data (features) and the best splitting of the nodes by minimizing the correlation between the trees. All supervised pixel-based classifications rely heavily on the input training samples. To discriminate croplands under various environments and condition, the sample size of the initial training dataset needs to be large, especially in complex regions. All samples were selected to represent a 90-m x 90-m polygon. First, we made extensive field campaigns in Australia during the 2014-crop growing season when data were collected on precise cropland locations as well as non-cropland locations. This effort led to collection of more than 628 samples spread across Australia (e.g., Table 3). Second, we absorbed the ground data from previous efforts for China and other reliable sources. Third, sub-meter to 5-m very high spatial resolution imagery, available for us for the entire study region, was used to generate croplands *versus* non-cropland interpretations by multiple analyses across China, Australia, and New Zealand and a total of ~1490 data samples were used from these interpretations. To move forward with a larger sample size, an iterative sample selection procedure was introduced with the following steps for training the Random Forest (RF) machine-learning algorithm as illustrated in Figure 1.

1. Build Random Forest classifier using existing training samples. Initially we start with a small number of samples and slowly increase the sample size till we reach high degree of accuracy and the accuracy plateaus at certain sample size;
2. Based on established classifier, classify 30-m MFDC using Random Forest algorithm in GEE cloud;
3. Visual assessment of classification results are compared with existing reference maps as well as sub-meter to 5-m very high spatial resolution imagery (VHRI); The process (Figure 1) was iterated until sufficient correspondence is achieved;
4. Added (see Figure 1) 'crop' samples in missing area and 'non-crop' samples by referencing sub-meter to 5-m very high spatial imagery from Google Earth Imagery. For cases hard to tell by interpretation (fallow-land or abandoned fields), historical Landsat Images and MODIS NDVI time-series are also referenced. All the samples selected to represent a 90-m x 90-m polygon.
5. Loop step 1-4 with enlarged training dataset until classification becomes stable.

The number of iterations required for the training sample selection is a function of the complexity of the area. China was divided into three zones; New Zealand was divided into two zones to carry out classification (Figure 2): the iterative selection will have to loop ~4-5 times to improve the initial classified results.

ii. Programming and codes

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

iii. Results

The machine learning algorithms (RF), discussed in previous sections, were trained to separate croplands *versus* non-croplands for each of the zones (Figure 2) based on knowledge generated using reference data. The machine learning algorithms were then run on the Google Earth Engine (GEE) cloud-computing environment using a Landsat-8 collection for each of the zones to separate croplands *versus* non-croplands. The process was iterated and knowledge in the algorithms tweaked several times, before getting accurate results of croplands *versus* non-croplands. This process led to producing the Global Food Security-support Analysis Data @ 30-m cropland extent for China (Figure 5), Australia (Figure 6), and New Zealand (Figure 7) product. This product is available through the Land Processes Distributed Active Archive Center (LP DAAC). The same dataset is also available for visualization at https://croplands.org/app/map.

Zoom-in views show complete resolution of the imagery that shows individual farms (Figure 5, 6, 7, and 8). Full resolution of 30-m cropland extent can be visualized in croplands.org by zooming-in to specific areas as illustrated in right panel (b) and (c) of Figures 5, 6, 7, and 8. For any area in Australia, China, New Zealand, or Mongolia croplands can be visualized by zooming into specific areas in croplands.org. The background sub-meter to 5-m imagery, available for the regions on Google Earth helps evaluate the precision of the cropland extent product (“zoom in” and “toggle” cropland “on” and “off” to see the sub-meter to 5-m imagery in the background).

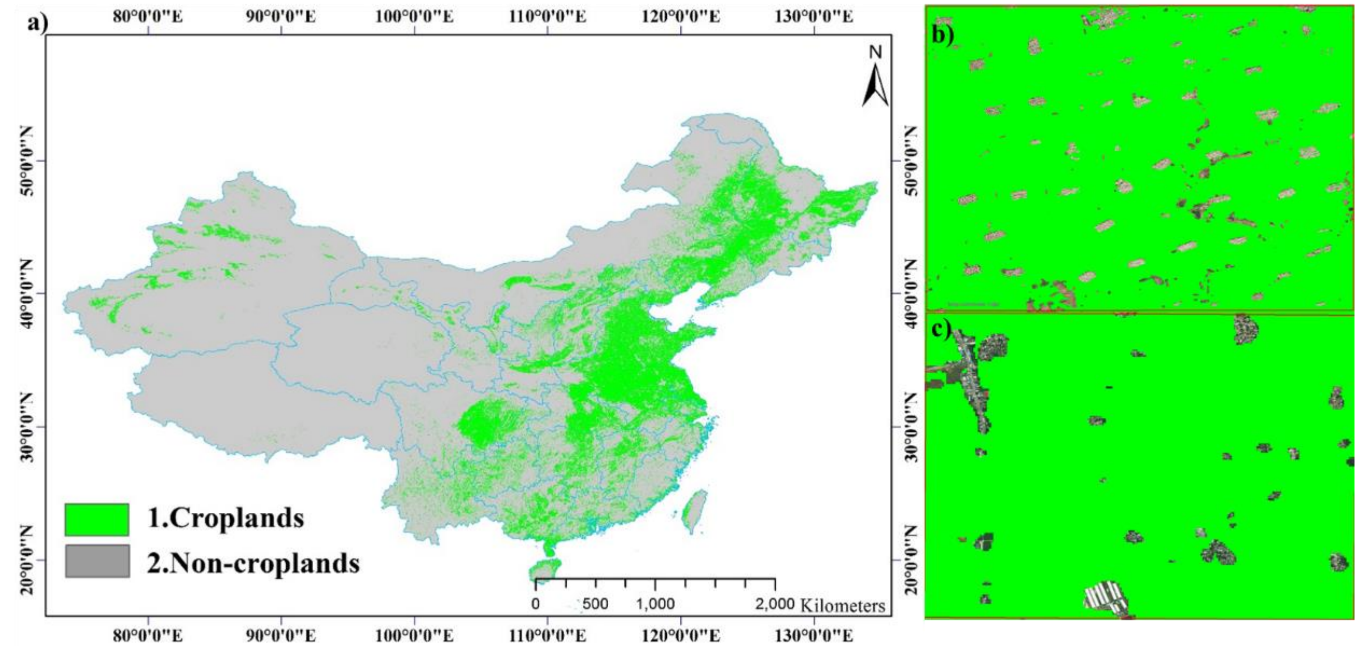

Figure 5. Cropland Extent Product at 30-m for China (left image) with illustrative zoom in view for a location (right). This product is made available for visualization @: croplands.org. The data are downloadable from LP DAAC.

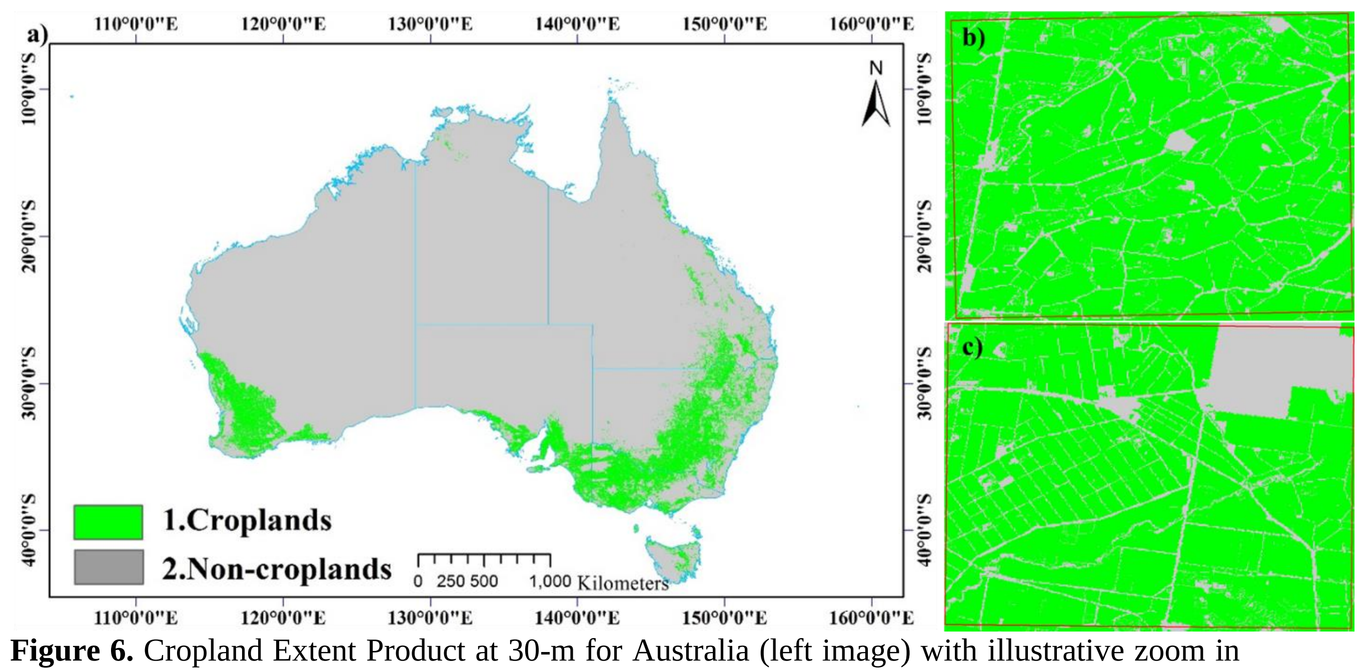

Figure 6. Cropland Extent Product at 30-m for Australia (left image) with illustrative zoom in view for a location (right). This product is made available for visualization @: croplands.org. The data are downloadable from LP DAAC.

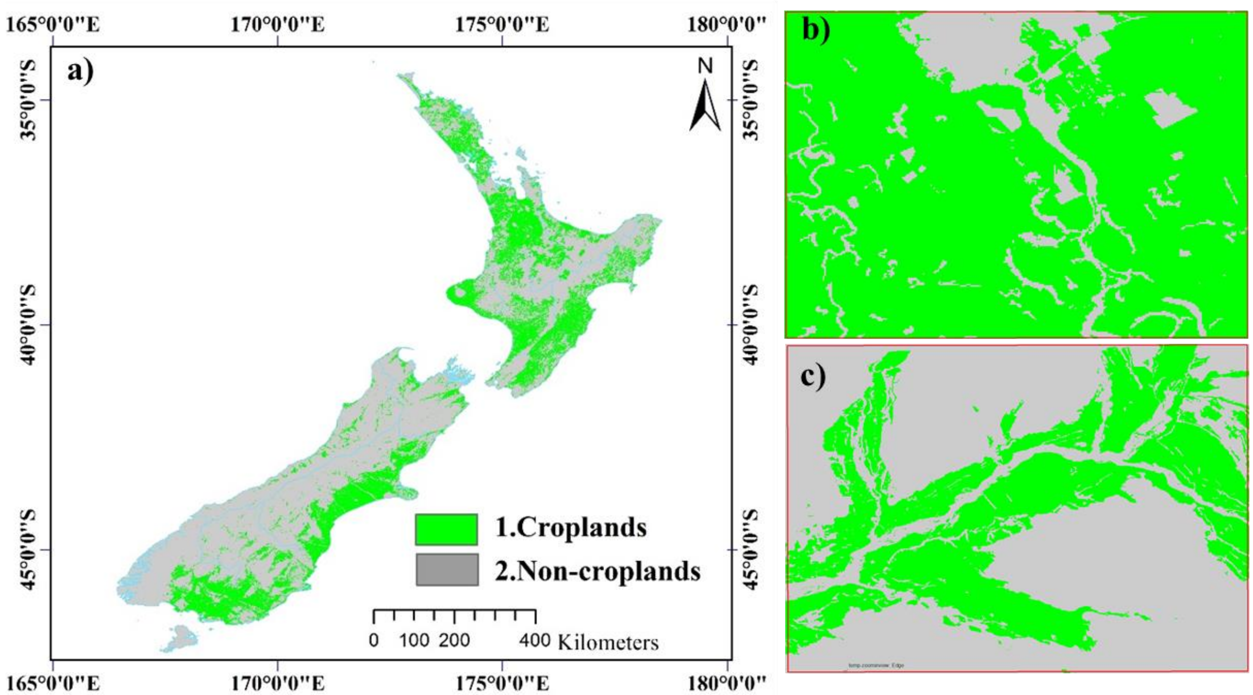

Figure 7. Cropland Extent Product at 30-m for New Zealand (left image) with illustrative zoom in view for a location (right). This product is made available for visualization @: croplands.org. The data are downloadable from LP DAAC.

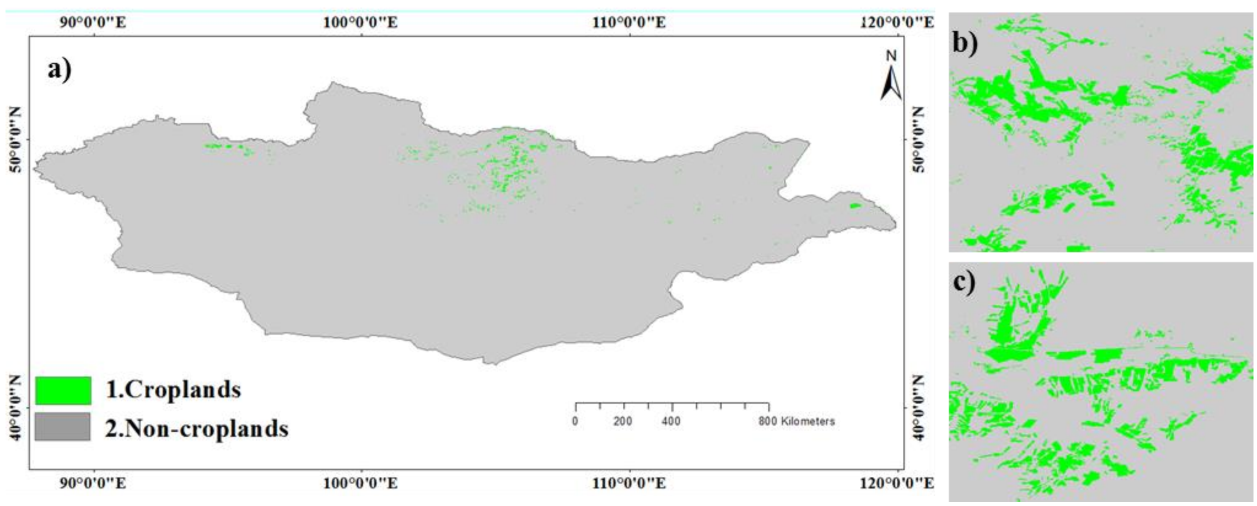

Figure 8. Cropland Extent Product at 30-m for Mongolia (left image) with illustrative zoom in view for a location (right). This product is made available for visualization @: croplands.org. The data are downloadable from LP DAAC.

iv. Cropland Areas

Country wise cropland areas calculated based on 30-m crop extent maps from this study are summarized here. Table 5 shows country-by-country cropland area statistics of all countries generated from this study and compared with several other sources such as the national census data based MIRCA2000 (Stefan Siebert and Portmann, personal communication; Portmann, 2010) which were also updated in the year 2015, The Food and Agricultural Organization (FAO) of United Nation's compiled statistics, MODIS 500-m derived cropland areas from GRIPC (Salmon et al., 2015), and GIAM-GMRCA (Thenkabail et al., 2009 and Biradar et al., 2009) derived cropland areas. Overall as per GFSAD30AUNZCNMOCE estimates, total net cropland areas of China, Australia, and New Zealand were 165.2 Mha, 35.1Mha and 2.25 Mha respectively. Further, the cropland areas of China were computed at sub-national level for the 31 provinces of China and were compared with cropland areas from national statistics of China (http://www.stats.gov.cn/tjsj/ndsj/2014/indexeh.htm) and shown in Figure 9. Among 31 provinces, two provinces show under-estimation and two provinces shows over-estimation relative to national statistics.

Table 5. Net cropland areas (NCAs) derived based on 30-m GFSAD30 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]**
Name	**Ha**	**Ha**	**Ha**	**Ha**	**Ha**	**Ha**
China	932,824,512	165,228,334	158,872,013	248,526,732	203,624,473	203,607,871
Australia	768,851,504	35,105,792	30,615,114	47,447,364	48,623,546	54,933,291
New Zealand	26,353,211	2,249,700	1,021,298	540,030	1,585,089	2,607,923
Mongolia	155,436,242	1,211,415	1,961,351	926,400	2,559,316	4,889,473

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

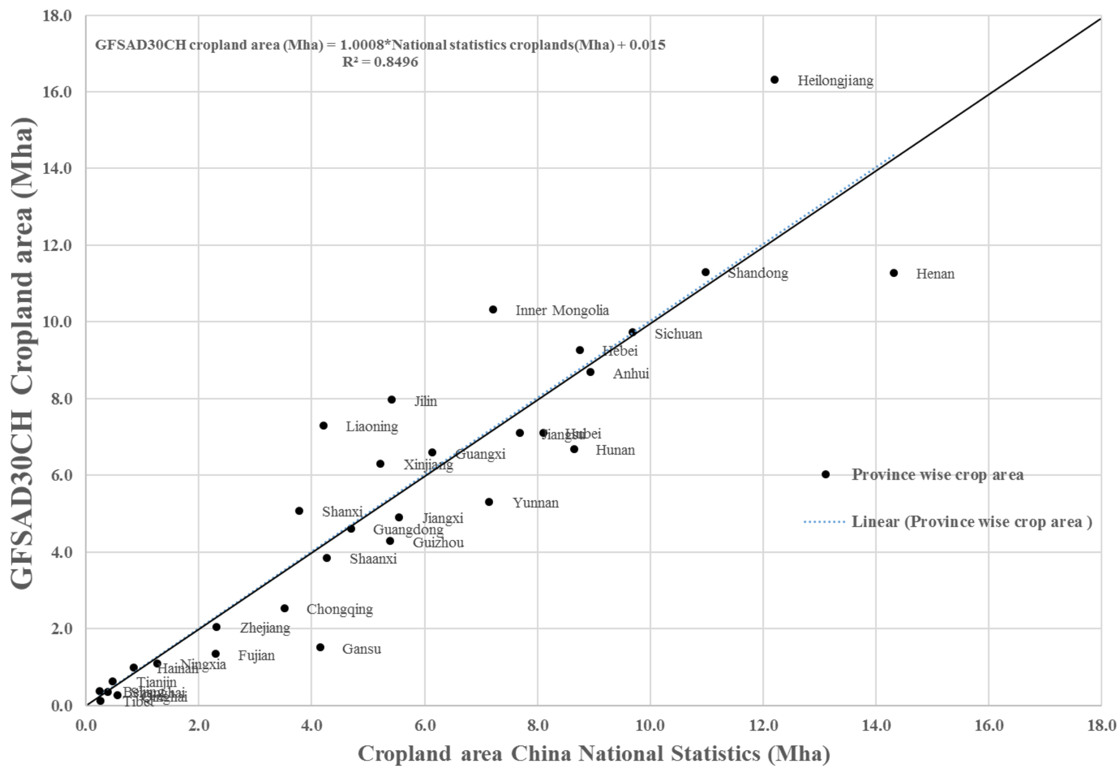

Figure 9. Comparison of province wise cropland areas derived from GFSAD30AUNZCNMOCE for China vs. cropland areas obtained from National statistics of China (http://www.stats.gov.cn).

V. Calibration Needs/Validation Activities

An independent accuracy assessment was performed zone by zone.

Australia: For the accuracy assessment in Australia, a total of 900 validation samples were used to determine the accuracies of the final cropland extent map of Australia. An error matrix (Table 6) was generated for all of Australia that provided producer's, user's, and overall accuracies (Story and Congalton, 1986, Congalton, 1991, and Congalton and Green, 2009). For all of Australia (which was in single zone# 68), the overall accuracies were 97.6% with producer's accuracy of 98.8% (errors of omissions of 1.2%) and user's accuracy of 79.0% (errors of commissions 21%) for the cropland class (Table 6). These results clearly imply the high level of confidence in differentiating croplands from non-croplands for Australia. High producer's accuracies for the cropland class across Australia (98.8%) suggested that almost all croplands of Australia were mapped accurately- with only 1.2% of croplands omitted. However, as a trade-off, for achieving such high producer's accuracies for the cropland class, there was an commission error of 21% (meaning non-croplands mapped as croplands). Since our goal is to map almost all croplands, this small compromise of mapping some non-croplands as croplands need to be accepted.

Table 6. Independent Accuracy Assessment error matrix of 30-m Cropland Extent Product for Australia.

Zone 68		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	79	21	100	79.0%
	No-Crop	1	799	800	99.9%
Total		80	820	900	
Producer Accuracy		98.8%	97.4%		**97.6%**

New Zealand: For accuracy assessment in New Zealand, 500 validation samples, combined for two zones in New Zealand, were used to determine the accuracies of the final cropland extent map of New Zealand. Error matrices (Table 7) were generated for each of the zones separately and for all of New Zealand providing producer's, user's and overall accuracies. For all of New Zealand, the overall accuracies were 93.4% with producer's accuracy of 91.7% (errors of omissions of 8.3%) and user's accuracy of 82.7% (errors of commissions of 17.3%) for the cropland class (Table 7). When considering both zones, the overall accuracies ranged between 90-96.8%, producer's accuracies of the cropland class in the Zone #69 and #70 were 92.3% and 91.4% respectively (Table 7), and user's accuracies for the cropland class in the Zone #69 and #70 were 87.8% and 80.4% respectively (Table 7).

Table 7. Independent Accuracy Assessment error matrix of 30-m Cropland Extent Product for New Zealand by Zone.

Zone 69		Reference Data			
% of TCA*=	18.8%	Crop	No-Crop	Total	User Accuracy
Map Data	Crop	36	5	41	87.8%
	No-Crop	3	206	209	98.6%
Total		39	211	250	
Producer Accuracy		92.3%	97.6%		**96.8%**

Zone 70		Reference Data			
% of TCA*=	81.2%	Crop	No-Crop	Total	User Accuracy
Map Data	Crop	74	18	92	80.4%
	No-Crop	7	151	158	95.6%
Total		81	169	250	
Producer Accuracy		91.4%	89.4%		**90.0%**

		Reference Data			
Combined		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	110	23	133	82.7%
	No-Crop	10	357	367	97.3%
Total		120	380	500	
Producer Accuracy		91.7%	94.0%		**93.4%**

TCA* (Total Cropland Area) = 8.1 Mha

China: For the accuracy assessment in China, a total of 1972 validation samples, combined for all three zones in China, were used to determine the accuracy of the final cropland extent map of China. Error matrices (Table 8) were generated for each of the zones separately and for all of China providing producer's, user's, and overall accuracies (Story and Congalton, 1986, Congalton, 1991, and Congalton and Green, 2009). For all of China, the overall accuracies were 94% with producer's accuracy of 80% (errors of omissions of 20%) and user's accuracy of 84.2% (errors of commissions of 15.8%) for the cropland class (Table 8). When considering all three zones, 95% of total cropland area of China is in zone #58 in Table 5, 4% of total cropland area of China is in zone #59, and only <1% of total cropland area of China is in zone #60. The overall accuracies ranged between 91-98%; producer's accuracies of the cropland class for the zone #58 and zone #59 were 83.3% and 65% respectively. Zone #60 had low overall accuracies, but it has only about 1% of the total cropland areas of China and hence is inconsequential for most applications. Zone 58, which has 95% of all China's croplands, had for the cropland class producer's accuracy of 83.3% and user's accuracy of 84.2% for the cropland class. This means, in this zone, 16.7% of the croplands were not mapped (omissions errors) and 15.8% of the non-croplands were mapped as croplands. These results clearly imply the high level of confidence in differentiating croplands from non-croplands for China. High producer's accuracies across zones suggest that few croplands were omitted during the mapping process.

Table 8. Independent accuracy assessment error matrix for the 30-m Cropland Extent Product for China by zone.

Zone 58		Reference Data			
% of TCA*= 95.1%		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	255	48	303	84.2%
	No-Crop	51	830	881	94.2%
Total		306	878	1,184	
Producer Accuracy		83.3%	94.5%		**91.6%**

Zone 60		Reference Data			
% of TCA*= 1.0%		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	4	0	4	100.0%
	No-Crop	10	335	345	97.1%
Total		14	335	349	
Producer Accuracy		28.6%	100.0%		**97.1%**

Zone 59		Reference Data			
% of TCA*= 3.9%		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	13	3	16	81.3%
	No-Crop	7	416	423	98.4%
Total		20	419	439	
Producer Accuracy		65.0%	99.3%		**97.7%**

Combined		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	272	51	323	84.2%
	No-Crop	68	1,581	1,649	95.9%
Total		340	1,632	1,972	
Producer Accuracy		80.0 %	96.9 %		**94.0%**

TCA* (Total Cropland Area) = 165 Mha

Mongolia: For the accuracy assessment in Mongolia, a total of 310 validation samples were used to determine the accuracies of the final cropland extent map of Mongolia. An error matrix (Table 9) was generated for all of Mongolia that provided producer's, user's, and overall accuracies (Story and Congalton, 1986, Congalton, 1991, and Congalton and Green, 2009). For all of Mongola (which was in single zone# 57), the overall accuracies were 98.4% with producer's accuracy of 75.0% (errors of omissions of 25%) and user's accuracy of 92.3% (errors of commissions 7.7%) for the cropland class (Table 8). These results clearly imply the high level of confidence in differentiating croplands from non-croplands for Mongolia.

Table 9. Independent accuracy assessment error matrix for the 30-m Cropland Extent Product for Mongolia.

Zone 57		Reference Data			
		Crop	No-Crop	Total	User Accuracy
Map Data	Crop	12	1	13	92.3%
	No-Crop	4	293	297	98.7%
Total		16	294	310	
Producer Accuracy		75.0%	99.7%		**98.4%**

VI. Constraints and Limitations

Note overlapping tiles: The following tile also covers part of another tile in GFSAD30SEACE (Indonesia). Please ignore the Indonesian data in the following tile and keep only Australian part.
GFSAD30AUNZCNMOCE_2015_S20E120_001_2017286154500.tif

GFSAD30AUNZCNMOCE product mapped the croplands of Australia, New Zealand, and China @ nominal 30-m, which is the best known resolution for cropland mapping over such a large Agriculture area. It also has high levels of overall accuracies of 97.6%, 93.4%, and 94.0 % for Australia, New Zealand, and China respectively, producer's accuracies of 98.8%, 913% 80.0%, and user's accuracies of 78.8%, 82.7%, and 84.4 % for cropland class.
A producer's accuracy of 98.8% for the cropland class of Australia means an error of omission of 1.2%. This means 1.2% of the continental croplands were missing in the product. A user accuracy of 78.8% for the cropland class for Australia means there is an error of commission of 22.2%. This means, 22.2% of non-croplands are mapped as croplands. We tweaked the machine learning algorithms (section IV) to maximize capturing as much croplands as feasible automatically. In this process, some non-croplands are mapped as croplands as well. This is a preferred solution, in order to not miss croplands or to only miss them minimally. As a compromise, mapping some non-croplands as croplands becomes unavoidable. Similar interpretations can be made for New Zealand and China based on their error matrices (Table 7 for New Zealand and Table 8 for China).

Numerous issues cause uncertainties and limitations in cropland extent product. Some of these issues are discussed here. First, temporal coverage. The 16 day Landsat-8 and 16-day Landsat-7 coverage when put together, lead to substantial temporal coverage. Yet, if we look at Figure 3, we were only able to achieve seasonal (and not bi-weekly or monthly) cloud-free or near cloud-free mosaics of the entire study area. This is not surprising given such a large area involved and frequent cloud across the China. As a result, if we were to have daily coverage over an area (e.g., like MODIS) then it becomes feasible to have more frequent (e.g., monthly or bi-monthly composites) temporal coverage of the continent that will help advance cropland mapping at improved accuracies. Currently, even with two Landsat satellites, at best, we have two 4 images per month (compared to 30 images of MODIS when we consider daily day time coverage of MODIS). Second, there is a need for greater understanding of the Landsat-7 and Landsat-8 data on how well they are correlated and in efforts to achieve better harmonization of data from two different sensors. Third, is the limitation of the reference training and validation data. In this project, we already have large training and validation data compared to any previous work as described in various previous sections. Nevertheless, much wider and extensive field visits to different parts of the study regions will be helpful in better understanding of the issues involved and as a result better mapping. We had extensive field visits in Australia and China, but these data were mostly acquired one time. Greatest difficulties in cropland mapping in China were in mountainous agriculture (e.g., terrace agriculture), cropland fallows (e.g., whether a fallow is 1 year or 5-year or permanent). These and numerous other issues (e.g., implementing machine learning algorithms and uncertainties inherent in them) will continue to be there in cropland mapping over such large agriculture areas in china. Nevertheless advances made in this study is significant, especially in developing a nominal 30-m cropland extent of a large agriculture countries like China, Australia and New Zealand at very good accuracies.

VII. Publications

1. Peer-reviewed publications relevant to this study

Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M., Congalton, R., and Yadav, K. (2017). 30-m Cropland Extent and Areas of Australia, New Zealand, and China for the Year 2015 Derived using Landsat-8 Time-Series Data for three years (2013-2015) using Random Forest Algorithm on Google Earth Engine Cloud Platform. In preparation.

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

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

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

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

2. Peer-reviewed publications within GFSAD project

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

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

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

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

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

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

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

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

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

3. Web sites and Data portals:

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

4. Other relevant past publications prior to GFSAD project

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

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

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

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

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

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

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

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

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

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

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

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

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

5. Books and Book Chapters

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

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

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

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

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

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

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

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

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

VIII. Acknowledgements

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

IX. Contact Information

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

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

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

For the 30-m cropland extent product of Australia, New Zealand, and China, please contact:
Pardhasaradhi Teluguntla at pteluguntla@usgs.gov
Prasad S. Thenkabail at pthenkabail@usgs.gov
Jun Xiong at jxiong@usgs.gov

More details about the GFSAD30 project and products can be found at: **globalcroplands.org**

X. Citations

Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Congalton, R.G., Oliphant, A.J., Sankey, T., Poehnelt, J., Yadav, K., Massey, R., Phalke, A., 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 Australia, New Zealand, China, Mongolia 30 m V001* [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MEaSUREs/GFSAD/GFSAD30AUNZCNMOCE.001

XI. References

Biradar, C.M., Thenkabail, P.S., Noojipady, P., Li, Y., Dheeravath, V., Turral, H., Velpuri, M., Gumma, M.K., Gangalakunta, O.R.P., & Cai, X.L. 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, 114-129

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.

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

Congalton, R.G. 2015. Assessing positional and thematic accuracies of maps generated from remotely sensed data."Remote Sensing Handbook" three-volume set: Remotely Sensed Data Characterization, classification, and accuracies, Taylor and Francis Inc.\CRC Press, Boca Raton, London, New York. Pp. 800+. Pp. 625-662.
http://www.fao.org/faostat/en/#data/QC

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment.
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.
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.

Latifovic, R., Homer, C., Ressl, 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.

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.

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.

Portmann, F.T., Siebert, S., & Döll, P. 2010. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global biogeochemical cycles, 24

Roy, D.P., Wulder, M.A., Loveland, T.R., C E, W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z., 2014. Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment 145, 154–172.

Salmon, J.M., Friedl, M.A., Frolking, S., Wisser, D., & Douglas, E.M. 2015. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. International Journal of Applied Earth Observation and Geoinformation, 38, 321-334

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.

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

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., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., & Cai, X. 2009b. "Global irrigated area map GIAM, derived from remote sensing, for the end of the last millennium." International Journal of Remote Sensing, 30, 3679-3733

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.

Appendix I
Consolidated Figures