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 Southeast and Northeast Asia: Cropland Extent Product (GFSAD30SEACE)

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		Modification made according to USGS reviewer's comments

Contents

I. Members of the team 4
II. Historical Context and Background Information 4
III. Rationale for Development of the Algorithms 5
IV. Algorithm Description 5
a. Input data 6
i. Study Area Stratification 7
ii. Reference Samples 7
iii. Satellite Imagery: Landsat 7 and 8 8
b. Theoretical description 10
i. Definition of Croplands 10
ii. Pre-processing 11
iii. Random Forest (RF) Algorithm 11
iv. Post Processing Image Enhancements 11
c. Practical description 12
i. Pre-processing 12
ii. Random Forest (RF) Algorithm 12
iii. Post Processing Image Enhancements 13
iv. Programming and codes 14
v. Cropland areas of SE and NE Asia 14
V. Validation 17
VI. Constraints and Limitations 19
VII. Publications 19
VIII. Acknowledgements 23
IX. Contact Information 24
X. Citations 24
XI. References 24

I. Members of the team

This Global Food Security-support Analysis Data 30-m (GFSAD30) Cropland Extent-Product of Southeast and Northeast Asia (GFSAD30SEACE) was produced by the following team members. Their specific role is mentioned.

Mr. Adam Oliphant, Geographer, United States Geological Survey (USGS), led the GFSAD30SEACE product generation effort. Mr. Oliphant was instrumental in the design, coding, computing, analyzing, and synthesis of the Landsat-7 and 8-derived nominal 30-m GFSAD30SEACE cropland product of Southeast and Northeast Asia for the nominal year 2015. He was also instrumental in 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 GFSAD30SEACE product. He made a significant contribution in writing the manuscripts, ATBD, user documentation, and providing scientific guidance on the GFSAD30 project.

Dr. Jun Xiong, Research Scientist, Bay Area Environmental Research Institute (BAERI) at the United States Geological Survey (USGS), assisted in the coding and computing required for this project. He also assisted in writing the 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.

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 GFSAD30SEACE.

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. Richard Massey, PhD student at the Northern Arizona University, shared his expertise in cloud computing and contributed the pixel sieving algorithm and code used in image post-processing

II. Historical Context and Background Information

In order to maintain food and water security in the future, monitoring global croplands is a must. Unfortunately, currently available cropland products have major limitations. Current cropland maps feature disadvantages such as coarse resolution, difficulties in differentiating irrigated areas from rainfed areas, absence of precise spatial locations of croplands, an absence of crop types and cropping intensities, and the absence of a web/data portal for the distribution of these products. The goal of the Global Food Security Support Analysis Data @ 30-m (GFSAD30) project is to remove these limitations.

This algorithm theoretical basis document (ATBD) contains an account of, in particular, the GFSAD30 cropland extent product for Southeast and Northeast Asia (SE and NE Asia) (GFSAD30SEACE, Table 1), which includes SE Asia (ASEAN member states), NE Asia (Japan and North and South Korea), the Indonesian Archipelago, Melanesia, Micronesia, and Polynesia, produced using Landsat-7 and 8 time-series satellite sensor data.

Table 1. Basic information of the Global Food Security support-Analysis Data @ 30-m cropland extent product for Southeast and Northeast Asia (GFSAD30SEACE).

Product Name	Short Name	Spatial Resolution	Temporal Resolution
GFSAD 30-m Cropland Extent Product of SE and NE Asia	GFSAD30SEACE	30-m	nominal 2015

III. Rationale for Development of the Algorithms

Mapping cropland extent, as well as establishing location and area is essential for continuing food security. This knowledge is also essential for studying relationships with water, health, socio-economic, geo-political, environmental, and ecological factors (Thenkabail et al., 2010; Bruinsma 2011). The availability of accurate maps is also required for development of all higher level cropland products. These products include crop watering method, cropping intensities, cropland fallows, crop-type maps, and assessment of cropland and crop water productivity. A lack of certainty in the cropland extent map effects the certainty of all higher level cropland products. Unfortunately, reliable cropland extent maps that are of sufficiently high spatial resolution do not exist for the entire study area. The study area features many limitations, such as persistent cloud cover, aquaculture, and permanent plantations that are hard to distinguish from native forests.

Multitemporal classification of satellite imagery has become an important tool in land-use and land-cover change (LULC) science at the regional, national, continental, and global scale (Giri et al. 2003). Until recently, such analysis at the continental and global scales were restricted to course resolution imagery like Advanced Very-High-Resolution Radiometer (AVHRR) 1 km and Moderate-resolution Imaging Spectroradiometer (MODIS) 250m-500m (Gumma et al. 2016). Due to the expansion of parallel processing and huge data centers, it is now possible to create global classified maps using Landsat 30m imagery for Global Forest Cover, (Hansen et. al 2013) and GlobeLand30 (Chen et al. 2015). Although Chen's team has produced a global landcover map at 30 m resolution, it is not optimized for agricultural mapping and significant errors between GlobeLand30 and national agriculture maps exist (Brovelli et al. 2015). Recently, 30 m resolution cropland maps have been created for large portions of SE and NE Asia but they fail to accurately map the entire spatial extent included here (Cheng et al. 2016; Hurni et al. 2017; Kontgis et al. 2015).

This document describes the development of the 30-m Cropland Extent-Product of SE and NE Asia (GFSAD30SEACE). The approach involved a supervised Random Forest (RF) classifier to retrieve crop extent results from pixel-based classification to provide precise agriculture field boundaries at 30-m resolution.

IV. Algorithm Description

The methodology used in this project (Figure 1) is briefly described in this paragraph to provide an overview and presented in detail in subsequent sections of this ATBD document. The process (Figure 1) involved combining year 2013-2016 16-day time-series Landsat 7 and 8 30-m data (10 bands) along with SRTM 30-m data (slope and elevation). All available spectral bands were used as input variables in addition to three calculated spectral indices (Table 3). First, the data were pre-processed by cloud mask and gap-filling on Google Earth Engine (GEE, Gorelick et al., in press). Second, seasonal mosaics were created for three periods when many of the crops in SE and NE Asia are grown: Period 1 (January-April), Period 2: May-August, and Period 3: September-December. Additionally, the standard deviations for all Landsat 7 and 8 images acquired in 2015 for the 10 bands above were computed. The resulting 42-band mosaic was used as the base imagery for the classification as listed in Figure 1. Third, reference data were generated throughout SE and NE Asia to train the RF classifiers. A total of 7,849

reference samples were used for this purpose. Fourth, the trained RF classifiers were used to classify for cropland or non-cropland by utilizing the 42-band mosaic. Fifth, the composite cropland product of SE and NE Asia was evaluated for accuracy using 1,750 test samples. The process was iterated until adequate accuracies were attained. During this step, the validation data was only available to the accuracy assessment team. Finally, the GFSAD30SEACE product can be visualized on croplands.org and can be downloaded from NASA's LP DAAC.

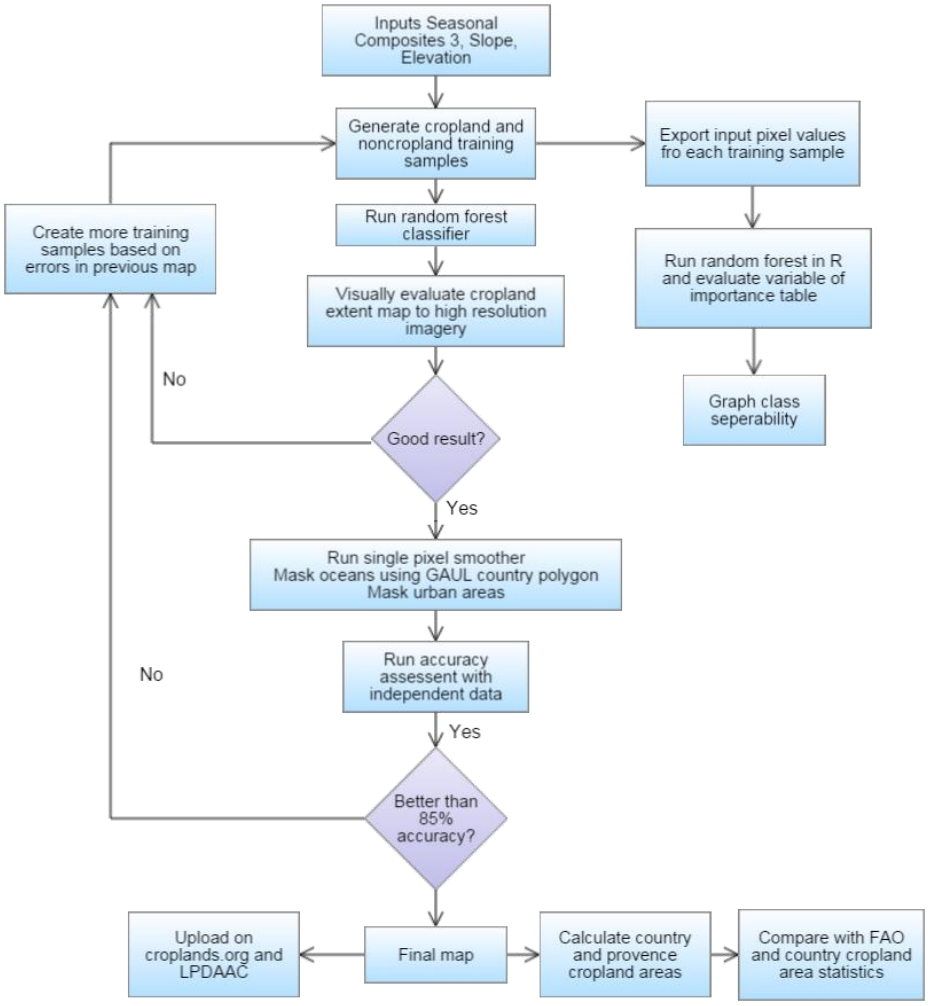

Figure 1. Flowchart of mapping methods for cropland extent product of SE and NE Asia utilizing Landsat 7 and 8 imagery for the nominal year 2015.

a. Input data

i. Study Area Stratification

GFSAD30 cropland extent product for Southeast and Northeast Asia (SE and NE Asia) (GFSAD30SEACE, Table 1), that includes SE Asia (ASEAN member states), NE Asia (Japan and North and South Korea), Indonesian Archipelago, Melanesia, Micronesia, and Polynesia. Since temperature regimes, precipitation patterns, crop calendars, and agricultural practices vary widely across SE and NE Asia, the area was broken into 7 refined agro-ecological zones (RAEZs) (Suepa et al 2016). These zones are based on FAO agro-ecological zones that were modified by the author to align with political boundaries and distinct agricultural areas. The 7 RAEZs are shown in Figure 2.

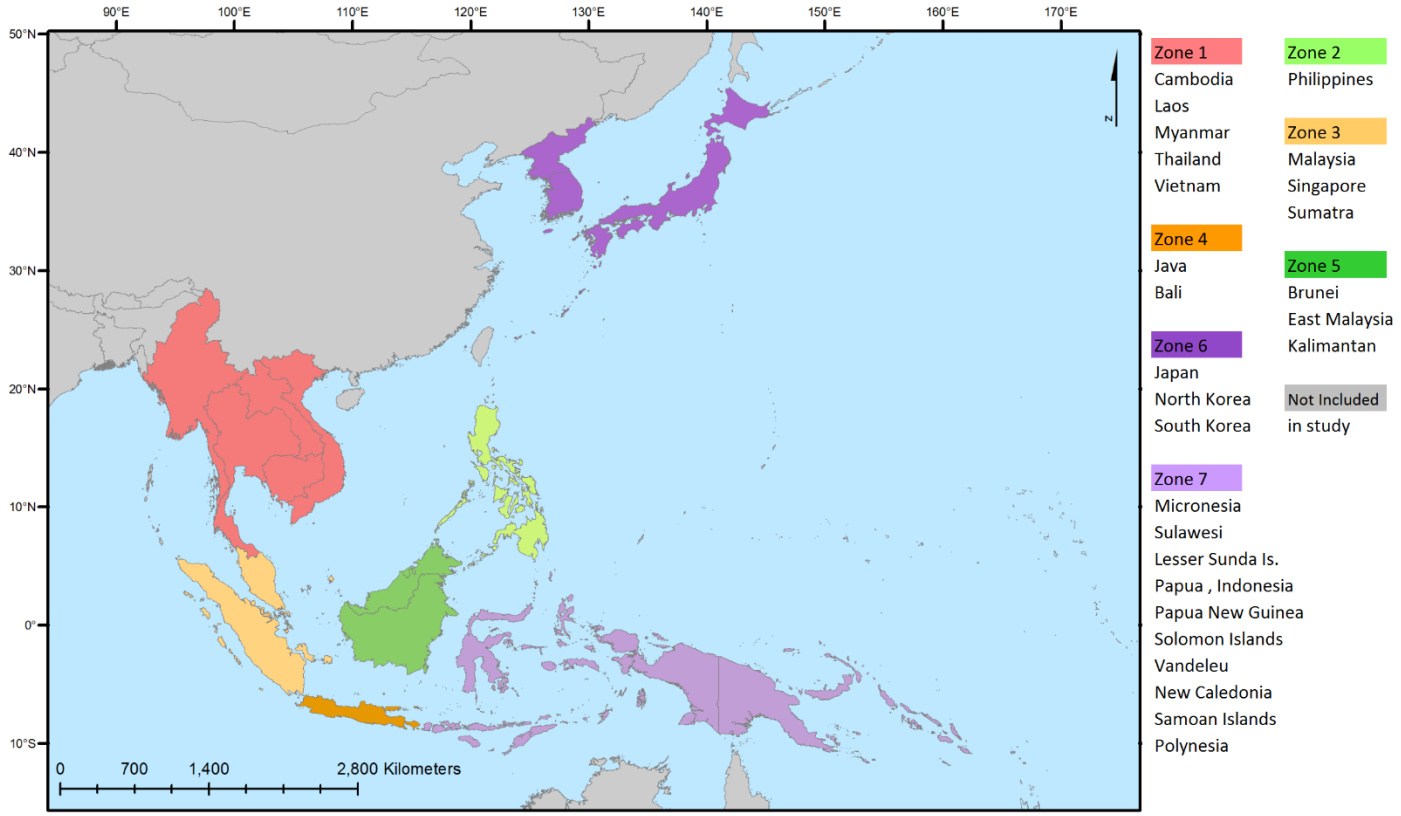

Figure 2. Southeast Asia study area and refined agro-ecological zones (RAEZs). Southeast Asia was broken into 7 zones for the pixel-based supervised Random Forest cropland extent classification. Zones were chosen based on temperature, seasonal precipitation, and farming practices. The zones are as follows: Zone 1 = Mainland SE Asia; Zone 2 = Philippines; Zone 3 = Sumatra and West Malaysia; Zone 4 = Java and Bali; Zone 5 = East Malaysia, Kalimantan, and Brunei; Zone 6 = Northeast Asia; Zone 7 = Pacific Island Nations.

ii. Reference Samples

Reference data that were used to train and validate the cropland product were collected by visually interpreting high spatial resolution imagery (VHRI). We generated the reference data by creating 400 random samples per zone, representing a 90 m x 90 m area. The samples were visually interpreted by referencing sub-meter to 5-meter very high spatial resolution imagery (VHRI) data throughout SE and NE Asia, made available to us from the National Geospatial Intelligence Agency (NGA). Samples were removed in areas where VHRI was not available and additional samples were placed in areas with poor classification results (Oliphant et al. 2017). There were a total 7,849 samples from VHRI spread across SE and NE Asia (see table 2). The spatial distribution is shown in Figure 3. The data used is available on at the following web site: **https://croplands.org/app/data/search**.

Table 2. Reference training data for each of the 7 refined agro-ecological zones (RAEZ's).

Zone	Region	Cropland Training Samples	Non Cropland Training Samples	Total Training Samples	Land Area	Land Area per Sample
		#	#	#	km^2	km^2 /#
1	Mainland SE Asia	1,326	1,267	2,593	1,939,900	748
2	Philippines	350	330	680	300,000	441
3	Sumatra & Malaysia	305	317	622	604,100	971
4	Java & Bali	298	272	570	134,100	235
5	Kalimantan	257	349	606	743,300	1227
6	Japan & Korea	460	481	941	598,700	636
7	Pacific Island Nations	614	1,223	1,837	1,253,800	683
	Total	3,610	4,239	7,849	5,573,900	710

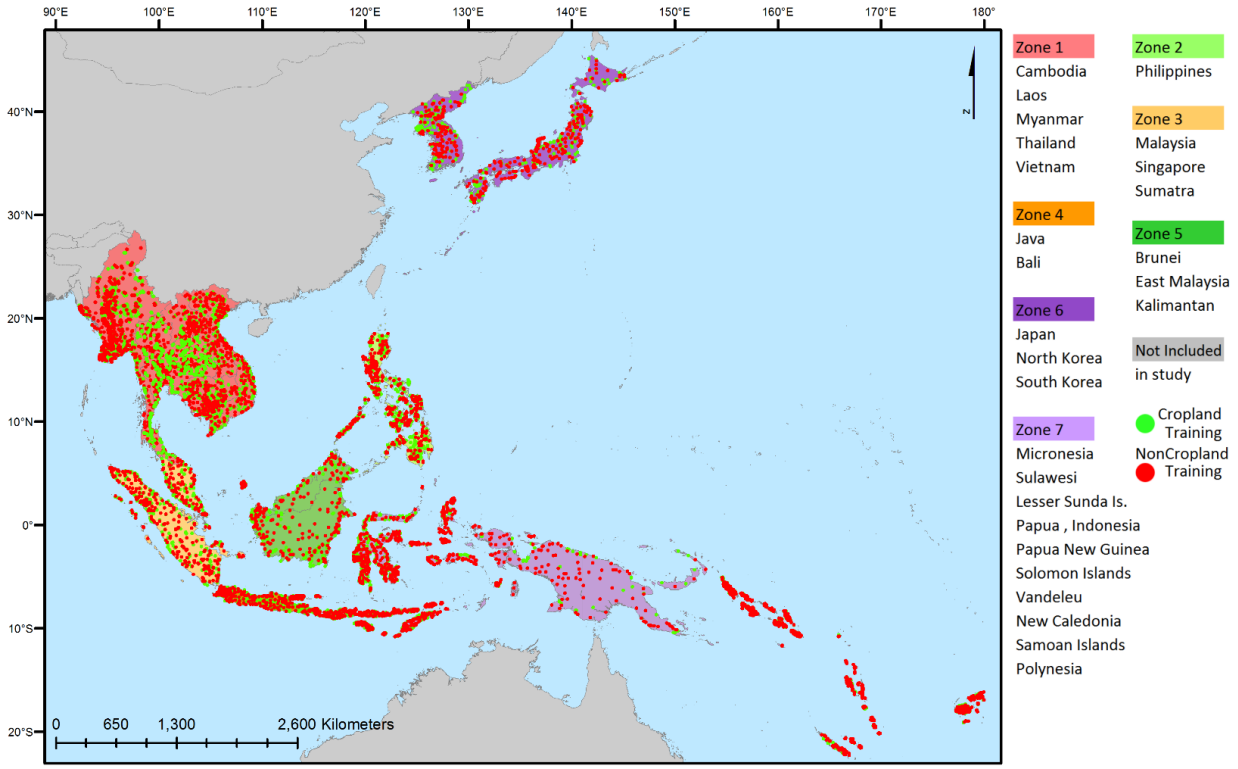

Figure 3. Classification zoning, training and validation samples of cropland extent product

iii. Satellite Imagery: Landsat 7 and 8

Landsat 7's Enhanced Thematic Mapper Plus (ETM+) and Landsat 8's Operational Land Imager (OLI) sensors are very useful in landcover mapping because the imagery is freely available and the sensors have excellent radiometric and spectral calibration (Roy et al. 2014). Landsat 7 and 8 each have a 16-day revisit time, 8 days if both sensors are used. However, the effective revisit time is much longer due to cloud cover and failure of Landsat 7 scan line corrector (Li et al. 2017).

ETM+ and OLI bands corresponding to blue, green, red, near infrared IR, SWIR 1, SWIR 2, and one thermal band along with the vegetation indices NDVI, NBR2, and LSWI were used for this classification (Table 3). Due to stray light entering OLI's thermal IR sensor, only Band 10 centered at 10.9 µm was used for thermal analysis of Landsat 8. The NDVI (Normalized Difference Vegetation Index) was selected to help distinguish dense vegetation including forests. The NBR2 (Normalized Burn Index 2) was chosen to distinguish barren and urban lands from other land cover. The LSWI (Liquid Surface Water Index) was included to help separate rice paddy and other bodies from land cover (Kontgis et al. 2015). Landsat TOA (top of atmosphere) reflectance was used instead of surface reflectance due to the limited temporal availability of Landsat 7 and 8 surface reflectance imagery in GEE at the time of classification. Landsat TOA imagery was provided by USGS according to methodology defined in the USGS Landsat Guidebooks (USGS 2017; USGS 2016).

Table 3. Characteristics of Landsat 7 and 8 data used in the study along with band indices.

Band Name	Landsat 8 OLI Spectral Range µm	Landsat 7 ETM+ Spectral Range µm
Bue	0.452 - 0.512	0.45-0.52
Green	0.533 - 0.590	0.52-0.60
Red	0.636 - 0.673	0.63-0.69
NIR	0.851 - 0.879	0.77-0.90
SWIR 1	1.566 - 1.651	1.55-1.75
SWIR 2	2.107 - 2.294	2.09-2.35
Thermal	10.60 - 11.19	10.40-12.50

VI Name	Equation
NDVI	$\frac{NIR - Red}{NIR + Red}$
NBR 2	$\frac{SWIR\ 1 - SWIR\ 2}{SWIR\ 1 + SWIR\ 2}$
LSWI	$\frac{NIR - SWIR\ 1}{NIR + SWIR\ 1}$

Note: NIR = near infrared, SWIR = shortwave infrared, OLI = operational land imager, ETM+ = Enhanced Thematic Mapper plus, NDVI = normalized difference vegetation index, NBR = normalized burn ratio, LSWI = land surface water index.

Cloud-free image composites were generated for 3 periods DOY 1-120, 121-240, and 241-365. Periods were chosen that approximately aligned with the crop growing seasons, temperature and precipitation trends and could be uniformly applied across all of SE and NE Asia (World Bank 2016). To insure at least one cloudfree pixel per season, data from 2013-2016 were used to form composites. The period composites were calculated by taking the median value from all unmasked pixels within the period within 2013-2016. Additionally, in order to capture interannual variation, a standard deviation band was computed for each of the 10 bands for the year 2015 to target the classification as nominal 2015.

Topographic information was included as an input variable in addition to the spectral information. Elevation and slope bands derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model Version 3 (Farr, 2007) at one arc-sec (approximately 30-m) resolution were added to the 40 period composites. These 42 bands were combined into a Google Earth Engine (GEE) Image Collection object, which enabled the full archive for SE and NE Asia to be accessed without specifying individual scenes so classification algorithms such as the RF could be executed in GEE. Figure 4 shows the compositing of the 10 bands over 4 periods, and the inclusion of elevation and slope bands to produce the 42-band variable input dataset.

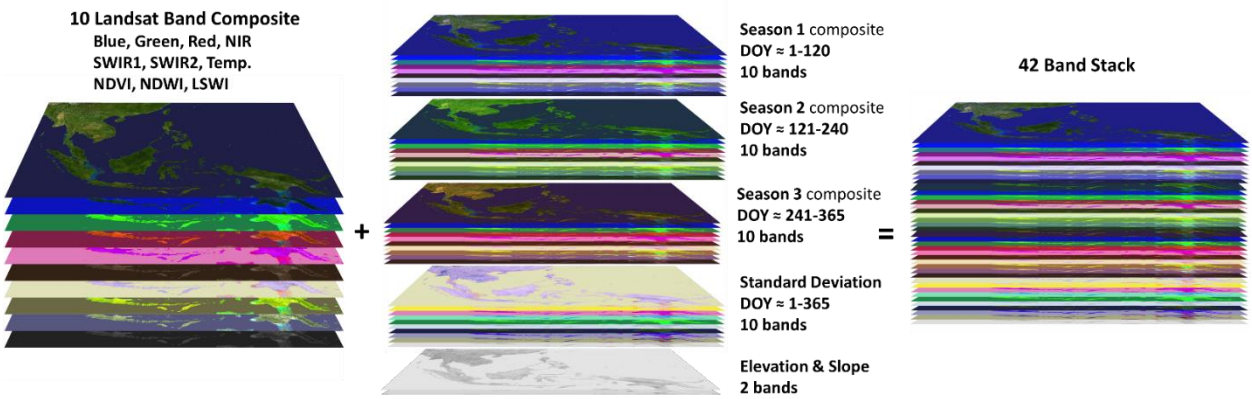

Figure 4. Data cube of 30-m Landsat 7 and 8 median value composites for 3 periods from years 2013 thru 2016. Also included are standard deviation for each image collected in 2015 along with STRM elevation and slope.

b. Theoretical description

i. Definition of Croplands

For all products within GFSAD30, cropland extent is 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 palm). 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 also includes cropland fallows. Non-croplands consist of all land cover classes other than croplands and cropland fallows (Figure 5).

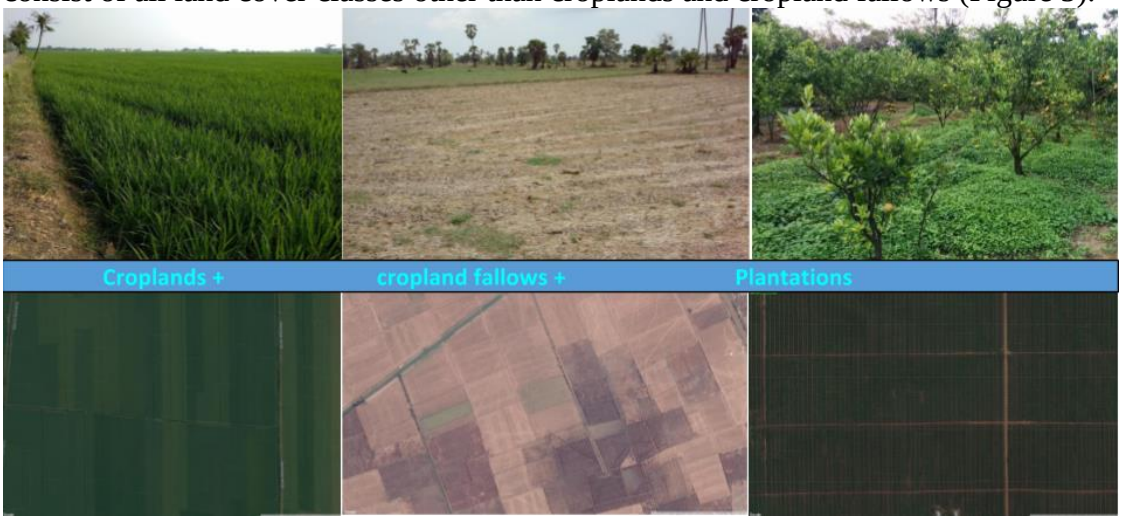

Figure 5. Illustration of definition of cropland mapping using aerial and ground images of crops. Croplands included: (a) standing crop, (b) cropland fallows, and (c) permanent plantation crops.

ii. Pre-processing

A cloud masking algorithm was applied as a pre-processing step prior to classification. The algorithm used was based on the Temporal Dark Outlier Mask, written by Carson Stam and Ian Housman was used to mask cloud and cloud shadow from Landsat 7 and 8 TOA imagery. The code is a refinement of the Landsat Simple Cloud Score algorithm which is a function inside GEE (Google Earth Engine API 2017).

The goal of this code is to identify clouds and mask out a sufficient radius around identified clouds to remove cloud shadows. Since this code does not seek to identify cloud shadows in particular, all pixels within a specific distance of the cloud are masked. This results in less data available to generate the composite and this was one of the reasons why composites needed to have multiple seasons and years of imagery.

iii. Random Forest

The random forest classifier is relatively fast, nonlinear classifier that excels in producing good results from noisy data (Rodriguez-Galiano et al., 2012; Pelletier et al., 2016). It uses bootstrap aggregating (bagging) to form an ensemble of trees by searching random subspaces from the given data (features) and the best splitting of the nodes by minimizing the correlation between the trees.

A random forest machine learning algorithm was used to do a binary classification of the SE and NE Asia region into cropland and non-cropland classes. Through experimentation, it was determined that 300 trees seemed a good balance between classification speed and accuracy. The default values were chosed for variablesPerSplit ($\sqrt{n_{bands}}$), minLeafPopulation (1), and bag Fraction (0.5). The outOfBagMode = true, in order to use different random subsamples of training data in generating trees in order to reduce model overfitting.

iv. Post-processing

Post-processing image enhancements were used to create the final cropland extent product. In pixel-based land cover classification projects, large homogenous areas of one landcover class can often contain single pixels or small numbers of pixels of another landcover class. Algorithms can be applied to remove these artifacts often called “salt and pepper,” improving the accuracy and the visual correctness of map products. Many image smoothing algorithms including focal filters (Eliason and McEwen 1990; Chan et al. 2005) were tried and produced better but still unsatisfactory results.

Although applying focal filters usually do produce a more visually appealing and more useful classified map, the process often distorts boundaries between classes, and actually make the map appear less accurate in these places. To overcome this shortcoming, a pixel-sieving algorithm, based on code written by Richard Massy (personal communication) was used for post classification. This code smoothed based on a larger window set by the user (section c. iii). Based on visual comparison with simple focal filters, the smoothing algorithm used was at effective at removing salt and pepper artifacts while creating fewer edge distortions.

An additional disadvantage of using post classification smoothers is that fine features such as roads are obscured in the smoothed product although they were present in the pre smoothed product. To correct for this, roads were masked out using Open Street Map Vector files. Additionally, coastlines and large waterbodies were given a separate classification to separate them from the non-cropland land class. The water masked used was all areas classified as ‘water bodies’ by the European Space Agency (ESA) GlobCover 2009, version 2.3 (Bicheron et al. 2011). In some cases, particularly in deltas, river edges, and rice paddies, areas classified as cropland were also identified as water in. In these cases, the cropland classification remained unchanged.

c. Practical description

i. Pre-processing

Composite periods were chosen based on crop calendars, crop growing seasons, temperature and precipitation trends (World Bank 2016). The goal was to generate as many composites as possible while insuring that the composited were temporally broad enough to contain cloud free pixels across the entire SE and NE Asia. Due to reduce complexity, composited were generated for same temporal periods across all RAEZs. Ultimately, cloudfree image composites were generated for 3 periods DOY 1-120, 121-240, 241-365.

Due to heavy cloud cover in the region, particularly in Kalimantan, Papua, and during monsoon season, there was not consistently one pixel that was not cloud masked when creating period composites for only 1 year. Untimely data from 2013-2016 was used to form the composites. The period composites were calculated by taking the median value from all unmasked pixels within the period within 2013-2016.

If no cloud free pixels were found within the set year and DOY range (which was occasionally the case in isolated areas in cities and rainforest), a composite generated over the year period for DOY 1-365 was used to fill these areas. This was done to insure that no data gaps remained which could negatively affect the classification.

The largest adjustable parameter in the cloud masking and compositing code other than temporal range is ‘cloudThreshold’. This is a parameter of the Landsat Simple Cloud Composite algorithm, which is the basis of the algorithm used. The algorithm created an additional band named ‘cloud’ which extimates the probability that the land shown in each pixel is obscured by clouds. The range is from 0- 100 with 0 having no chance of being a cloud and 100 being complete confidence that the pixel was obscured by coudcover. Recommended values range from 10 to 30. After experimentation and visual comparisons, a cloudThreshold value of 20 was chosen, as it appeared to mask the majority of stratus, cumulus, and altocumulus clouds without masking much cloud free pixels (personal observation).

ii. Random Forest (RF) Algorithm

Because the sample size of the initial training dataset needs to be large for the RF classifier to work in such a complex region, the sample size was increased using an iterative sample selection procedure for training as illustrated in Figure 2.

1. Start with a large and high quality set of training samples for the random forest classifier that cover the variability in landcover within one of the areas or RAEZs.
2. Extract the band values of the 42-band mosaic of pixels that are in the training dataset. Use this information to build a random forest classification model.
3. Use above model to classify the RAEZs using the 30 m, 42-band mosaic.
4. Visually assess the classification results to compare with existing reference maps as well as sub-meter to 5-m very high spatial resolution imagery (VHRI);
5. Add 'cropland' and 'non-cropland' samples in miss-classified areas by referencing VHRI. For cases hard to tell by interpretation (fallow land or abandoned fields), reference historical Landsat images. All the samples selected represent a 90 m x 90 m polygon.

6. Loop steps 1-5 with enlarged training dataset until classification becomes stable.

The number of iterations required to achieve satisfactory classification was related to the complexity of the area. In RAEZs with well-developed agricultural land and strong seasons like Japan and Korea, relatively few iterations were needed whereas in Kalimantan, similarities between native forest and palm oil cause a lot of confusion and require more iterations to produce satisfactory results.

iii. Post-processing Image Enhancements

The smoothing algorithm has two stages, filling small gaps in large areas of cropland, and removing small isolated cropland areas surrounded by large areas of non-cropland.

To remove small isolated areas of cropland (less than 12 Landsat pixels) surrounded by large areas on non-cropland, the binary cropland non-cropland image was inversed so 0= cropland and 1= noncropland. Then a focal filter was extended to remove cropland pixels that did not have at least 12 neighbors. This was accomplished by resampling the map to 15 m x 15 m pixels and removing all cropland pixels that had fewer than 50 neighbors. Resampling was done since results looked superior than 30 m resolution. 50 was chosen as the minimum pixel number because one 30 m pixel = four 15 m pixels (4 x 12 = 48).

To fill isolated gaps between large sections of cropland, a focal filter was extended to remove non-cropland pixels that did not have at least 25 non-cropland neighbors. This process was not sensitive to resampling, so the native 30 m pixels were used to test neighbors. The number 25 was chosen based on visually reviewing several test sites across SE and NE Asia in increments of 5 from 10 to 50.

Major roads were removed from the cropland extent layer. The most current version of Open Street Map (as of March 30 2017) was downloaded from https://extract.bbbike.org/ and minor roads (which could contain cropland) were removed. This vector file was then converted to a 30m raster file, where the centerline of the remaining road was masked, regardless of the road width. The following road classes as defined in the attribute table were retained (type = motorway or primary or secondary or truck). Figure 6 illustrates the effect of the two post processing steps to the final image.

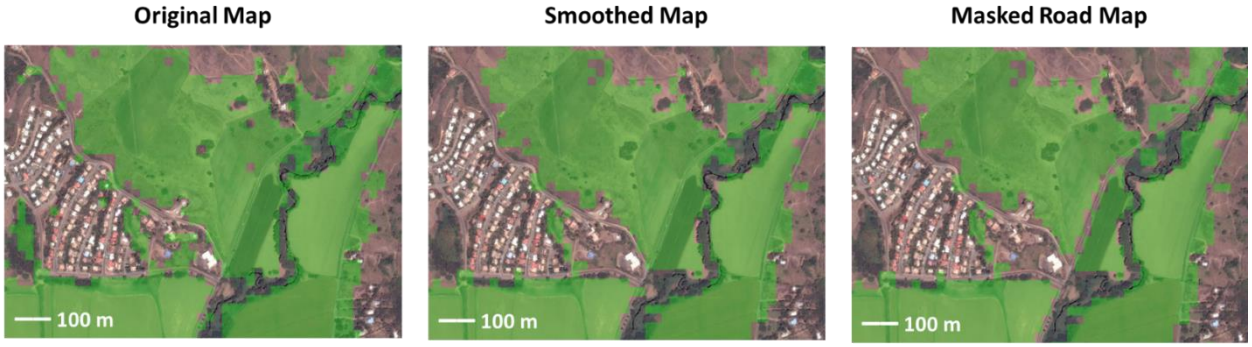

Figure 6. An illustration of the effects of pixel sieving and masking out major roads overlaid on VHRI where cropland is green and non-cropland is transparent. a) An image output from the stabilized RF classification. b) The same image after the pixel sieving algorithm has been applied. c) The same image after the major road mask has been applied. This image is centered at, 21.555°S, 165.494°E in New Caledonia.

A water mask was applied to improve the visual look and usefulness of the cropland extent product. This mask was applied such that, if a pixel was classified as both cropland and water, its final classification was cropland.

The mask used was the 'water bodies' class within the 2009 Globcover version 2.3 product, released December 2010 (Bicheron et al. 2011). Globcover is based on ESA MERIS satellite data and has a 300 m x300 m nominal ground resolution cell. This product was chosen because it has a global coverage, is rather recent (circa 2009) and has a 93% user's accuracy for the water bodies class (Bontemps et al. 2010).

iv. Programming and codes

The cloud masking and compositing algorithms and the supervised machine learning algorithm, random forest, were coded on Google Earth Engine (GEE) using its Application Programming Interface (API) with Java scripts. The codes are made available in a zip file and are available for download along with this ATBD.

v. Cropland areas of SE and NE Asia

The GFSAD30SEACE dataset 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**. Figure 7 shows the cropland extent in black overlaid the 7 RAEZs that were used to stratify SE and NE Asia. For any area in SE and NE Asia, croplands can be visualized by zooming into specific areas in **croplands.org.**

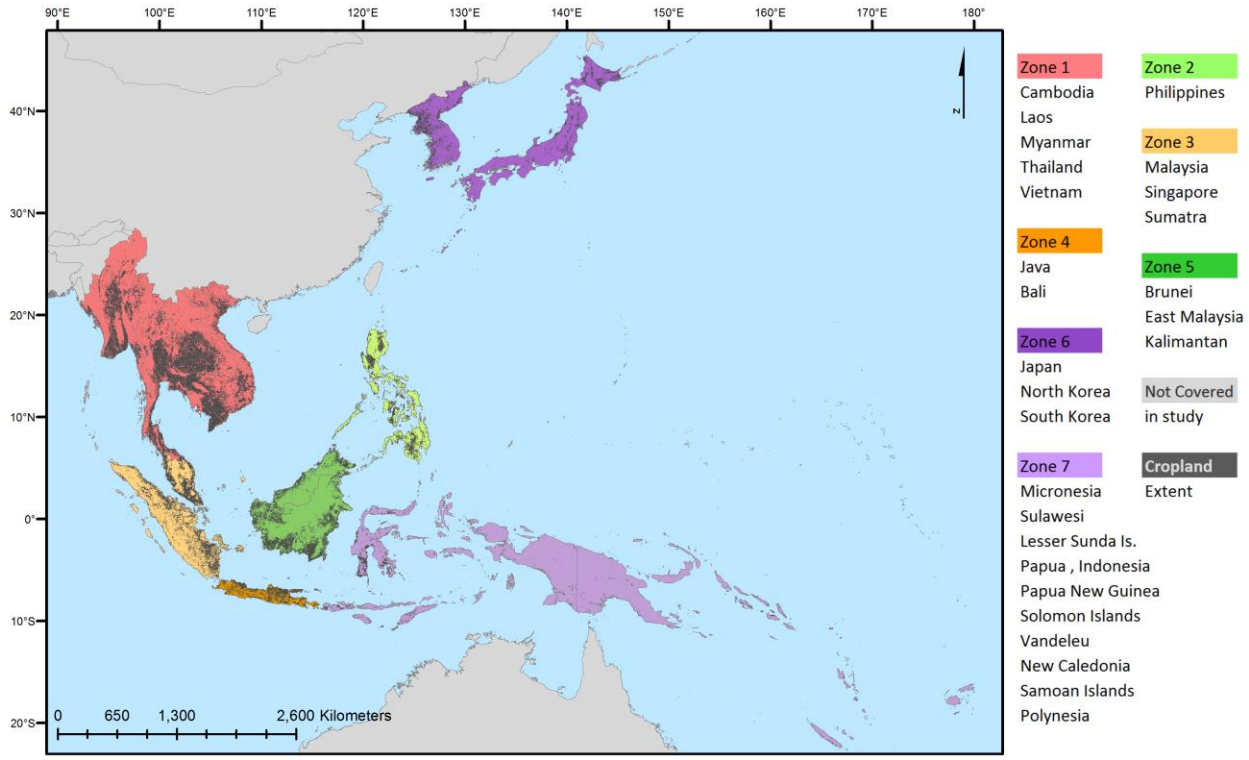

Figure 7. Cropland Extent Product at 30-m for Southeast andR Northeast Asia in black(GFSAD30SEACE)

Table 5 shows country-by-country cropland area statistics of SE and NE Asia, generated from this study and compared with several other sources. These sources included the MIRCA 2000, which is based on national census data (Stefan Siebert and Portmann, personal communication; Portmann, 2010) which were also updated in the year 2015, FAO of United Nation's compiled statistics (FAO 2013), MODIS 500-m derived cropland areas from GRIPC (Salmon et al., 2015), and GLC30 (Chen et al., 2015). SE and NE Asia had total net cropland area (TNCA) estimated from GFSAD30SEACE of 127,540 kha. All other studies estimated TNCA between 93,270 kha for GIAM to 135,260 kha for MIRCA. FAO-estimated cropland area was 135,260 kha for the year 2009. The overwhelming proportion of the cropland areas (90%) Table 5 were in just 7 of the 18 countries. Compared to

MIRCA2000 (Figure 8), the GFSAD30SEACE sometimes over-estimates and some-times under-estimates cropland areas.

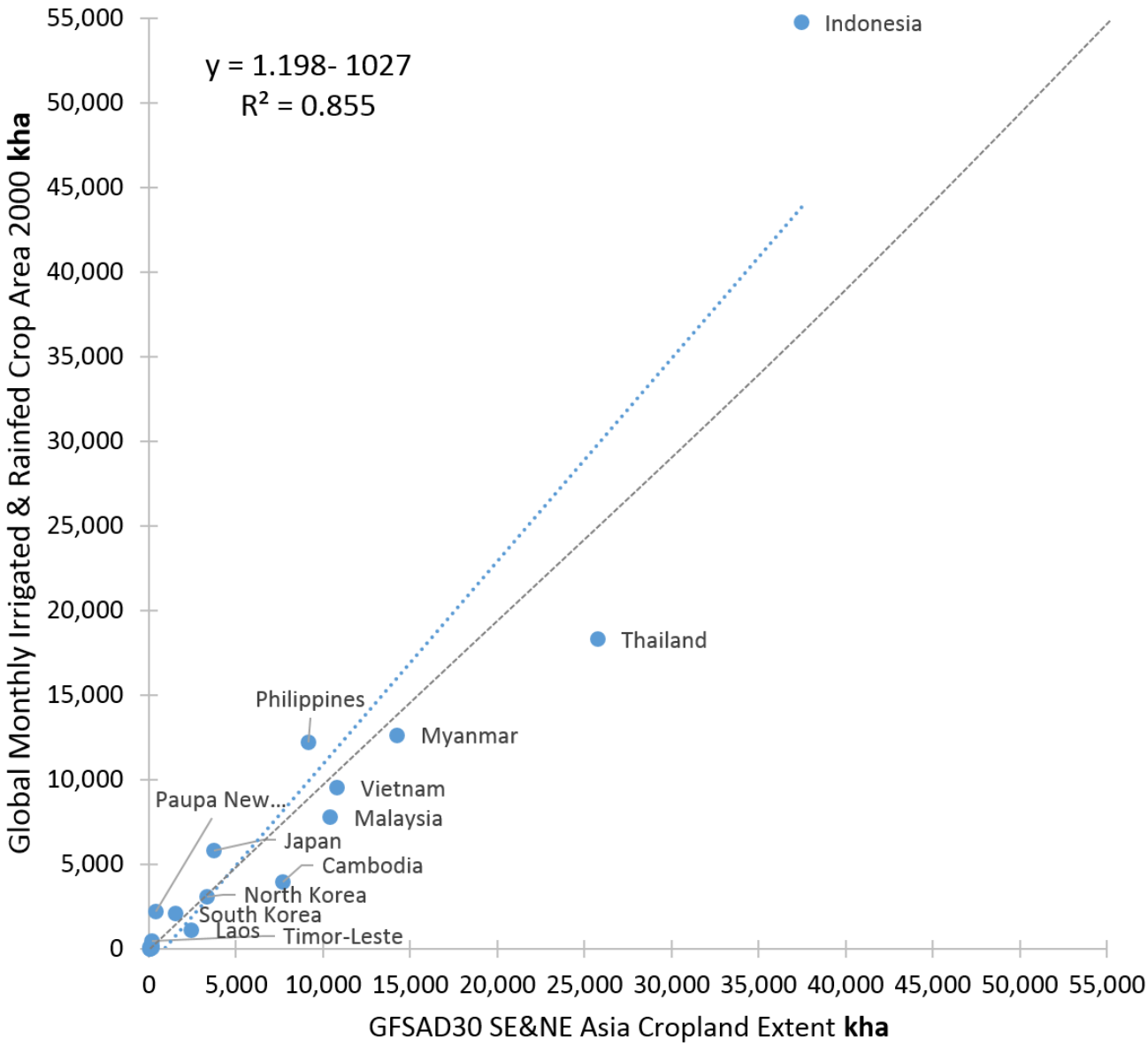

Figure 8. Country-by-country scatter plot of GFSAD30CESEA and MIRCA2000 (Portmann, 2010).

On average, the GFSAD30SEACE underestimated MIRCA by about 20% relative to statistical data according to the slope derived from Figure 8. This result is largely due to a large difference between the cropland area reported my GFSAD30CESEA and MIRCA2000 at 37,440 and 55,7500 kha respectively. However, the cropland area reported by MIRCA is likely high as FAO reports cropland as 42,600 kha. Other reasons for these significant differences in areas between GFSAD30SEACE when compared with MIRCA and FAO estimates include:

(i) Ability of 30-m derived croplands data of GFSAD30SEACE to capture fragmented croplands;
(ii) Ability of 30-m derived croplands data of GFSAD30SEACE to account for actual areas when compared with sub-pixel areas of higher resolution imagery derived cropland products (e.g., GRIPC, GLC);

(iii) Differences in how cropland data are gathered/estimated/calculated- Agricultural data reported by FAO is compiled from the statistics reported by individual countries, which is based on a wide range of methods, techniques and data (Bruinsma, J. 2011; World Bank, 2010). Methods include: questionnaire-based surveys of farmers and food distributors, subjective eye estimates of agricultural extension agents, independent evaluations from NGOs, best estimates from national officials, and mixed methods. Since each country uses different methods, accuracies are variable. Some countries may not have resources to maintain proper statistics, which can result in erroneous results. Errors in country-based statistics could be due to, resource limitations, data quality issues, and a lack of objectivity (political or economic motivations). The 30-m derived cropland data of GFSAD30CESEA provides objective estimates relative to how other statistical data are obtained.

(iv) Definition issues. Not every country adheres to the same definitions of croplands. Our study used TNCA definition to include planted crops along with croplands left fallow as well as permanent crops such as plantations (e.g., fruit trees, coffee and tea plantations, oil palm plantations etc). Many countries use similar definitions and many others use different definitions (e.g., leaving out cropland fallows or plantations).

(v) Uncertainties inherent in all estimates. One can expect uncertainties in cropland areas maps (e,g., Figure 7) or areas estimated from different sources (e.g., Table 4) as a result of definitions, data used, methods adopted, and reporting mechanisms. GFSAD30SEACE uncertainties are gauged by the error matrices (Table 4). For GFSAD30SEA, uncertainties in cropland estimates mainly arose from three sources: (a) permanent cropland including palm oil, (b) aquaculture, (c) greenhouses. In SE and NE Asia, large areas of native forest have been cut and oil palm plantations have been planted in their place. In areas where permanent crops including oil palm and rubber have fully matured, it is difficult to separate these crops from native forest. Aquaculture is an important part of the farming system in SE Asia, frequently occurring as shrimp farms near the coast and stock water fishponds near rice fields that serve as a source and sink for irrigation as well as providing farmers additional income and protein from the fish in them. However, some ponds are not used for aquaculture and uncertainty exist between aquaculture, rice paddy, and waterbodies exclusively used for irrigation. Per our definition, greenhouses are croplands, but difficult to map due to their differences in uncovered vegetation. Frequently, greenhouses are clustered together and can be identified visually with sub-meter to 5-m very high-resolution imagery. In such areas, greenhouses were usually classified as croplands. In areas where greenhouses are infrequent, they were often classified as noncropland.

There were several “outlier” countries in cropland area estimation when the GFSAD30SEACE product is compared to MIRCA2000 (Figure 8). There are many reasons for this. For example, GFSAD30SEACE methods and approaches are purely remote sensing based as opposed to predominantly survey-based statistics used in MIRCA2000. MIRCA2000 is a derived gridded dataset based on the FAOSTAT database (Portmann, 2010). Cropland was particularly hard to classify in many small Pacific Island Nations such as the in Solomon Islands and New Caledonia largely because many croplands are small fields surrounded by forest. In order to not omit agriculture, when composition of training samples was being modified, a greater emphasis was placed an adding cropland samples in omitted areas than adding non-cropland samples in co-mitted areas. This resulted in large amounts of natural vegetation and forest being classified as cropland in some areas. In other countries, it was difficult to distinguish permanent cropland from native forest so there were fewer cropland training samples than would be preferred, which resulted in lower cropland areas; some examples of this were in Papua New Guinea, Timor-Leste, and Vanuatu. Other countries have experienced rapid agricultural development in recent years, which is not reflected in previous studies; Cambodia is the best example of this.

Table 4. Net cropland areas of Southeast and Northeast Asia based on 30-m cropland product and comparison with other cropland products.

Country	Total Land Area (Kha) FAO-GAUL	Crop Area (Kha)				
		This Study:	Crop Ext. MIRCA	FAO Cultivated Area (2009)	GRIPC	GIAM
Resolution	--	30 m	variable	variable	500 m	1 km
Indonesia	181,100	37,440	54,750	42,600	31,030	20,750
Thailand	51,150	25,760	18,320	19,000	27,360	16,540
Myanmar	65,500	14,240	12,640	12,130	11,560	10,710
Vietnam	31,000	10,800	9,520	9,640	12,060	10,350
Malaysia	32,800	10,420	7,770	7,590	4,700	5,300
Philippines	29,800	9,150	12,190	10,440	12,820	9,020
Cambodia	17,600	7,680	3,980	4,060	4,980	4,600
Japan	36,580	3,730	5,790	4,610	5,990	5,950
North Korea	12,050	3,330	3,060	2,860	5,250	3,070
Laos	23,000	2,450	1,110	1,470	1,750	2,020
South Korea	9,700	1,520	2,090	1,780	3,180	3,120
Paupa New Guinea	46,000	370	2,220	960	1,690	1,610
Fiji	1,800	176	160	359	336	
Timor-Leste	1,500	178	451	225	0	226
Solomon Islands	2,800	156	78	76	20	0
Brunei	577	54	13	11		
Vanuatu	1,200	43	102	145	29	
New Caldonia	1,860	42	12	12	159	

V. Validation

For this assessment, 1,750 samples over the all 7 RAEZs in SE and NE Asia were used to determine the accuracy of the final cropland extent map of SE and NE Asia. Error matrices were generated for each of the zones separately and for the entire SE and NE Asia region providing producer's, user's and overall accuracies (Story and Congalton, 1986, Congalton, 1991, and Congalton and Green, 2009).

Table 4. Independent assessment of overall-, producer-, and user-accuracies of Croplands for 7 zones of Southeast and Northeast Asia. TNCA = Total Net Cropland Area of SE and NE Asia

Zone 1: Mainland SE Asia: 47.6 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**80**	20	100	20.0%
Non cropland	16	**134**	150	10.7%
Column total	96	154	**250**	
Omission error	16.7%	13.0%		
Producer accuracy	83.3%	87.0%		
User accuracy	80.0%	89.3%		
Overall accuracy				**85.6%**

Zone 2: Philippines: 7.2 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**41**	7	48	14.6%
Non cropland	15	**187**	202	7.4%
Column total	56	194	**250**	
Omission error	26.8%	3.6%		
Producer accuracy	73.2%	96.4%		
User accuracy	85.4%	92.6%		
Overall accuracy				**91.2%**

Zone 3: Sumatra & Malaysia: 14.5 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**45**	20	65	30.8%
Non cropland	22	**163**	185	11.9%
Column total	67	183	**250**	
Omission error	32.8%	10.9%		
Producer accuracy	67.2%	89.1%		
User accuracy	69.2%	88.1%		
Overall accuracy				**83.2%**

Zone 4: Java & Bali: 4.7 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**89**	27	116	23.3%
Non cropland	13	**121**	134	9.7%
Column total	102	148	**250**	
Omission error	12.7%	18.2%		
Producer accuracy	87.3%	81.8%		
User accuracy	76.7%	90.3%		
Overall accuracy				**84.0%**

Zone 5: Borneo: 14.4 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**69**	26	95	27.4%
Non cropland	2	**153**	155	1.3%
Column total	71	179	**250**	
Omission error	2.8%	14.5%		
Producer accuracy	97.2%	85.5%		
User accuracy	72.6%	98.7%		
Overall accuracy				**88.8%**

Zone 6: Japan & Korea: 6.7 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**42**	7	49	14.3%
Non cropland	15	**186**	201	7.5%
Column total	57	193	**250**	
Omission error	26.3%	3.6%		
Producer accuracy	73.7%	96.4%		
User accuracy	85.7%	92.5%		
Overall accuracy				**91.2%**

Zone 7: Pacific Island Nations: 4.9 % of TNCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**10**	7	17	41.2%
Non cropland	2	**231**	233	0.9%
Column total	12	238	**250**	
Omission error	16.7%	2.9%		
Producer accuracy	83.3%	97.1%		
User accuracy	58.8%	99.1%		
Overall accuracy				**96.4%**

Total: Entire Study Area: 100 % of TCA

	Cropland	Non cropland	Row total	Commission error
Cropland	**376**	114	490	23.3%
Non cropland	85	**1175**	1260	6.7%
Column total	461	1289	**1750**	
Omission error	18.4%	8.8%		
Producer accuracy	81.6%	91.2%		
User accuracy	76.7%	93.3%		
Overall accuracy				**88.6%**

The independent accuracy assessment team systematically tested each of the 7 refined agro-ecological zones or RAEZs (Figure 2) for accuracies. For the entirety of SE and NE Asia, the overall accuracy was 88.6% with producer's accuracy of 81.6% (errors of omissions of 18.4%) and user's accuracy of 76.7% (errors of commissions of 23.3%) for the cropland class. For each of the 7 RAEZs individually, the range of the (Table 4): (a) overall accuracies were 83.2-96.4%, (b) producer's accuracies were 67.2-97.2%, and (c) user's accuracies were 58.8-85.7%.

The overall accuracy in each zone ranged from 83% to 96%. Zone 7 had the highest overall accuracy since it has a relatively low proportion of agricultural area, which increased the overall accuracy. Overall, users' accuracy (commission errors) was relatively lower than producer's accuracies (omission errors) in general, which is the expected result when the RF algorithm is optimized to include as much croplands as much as possible so as to include all agricultural areas and permanent cropland. This results in a slight over classification of cropland area which lowers user's accuracy.

VI. Constraints and Limitations

GFSAD30SEACE product mapped the croplands of SE and NE Asia @ nominal 30-m, which is the best known resolution for cropland mapping over all of Southeast Asia, Northeast Asia and Pacific Islands. It also has high levels of accuracies with weighted overall accuracies of 88.6%, producer's accuracy of 81.6% and user's accuracy of 76.7%.

We tweaked the machine learning algorithms (section IV) to maximize capturing as much croplands as feasible automatically. This is a preferred solution to minimize how many croplands are missed/omitted from the final maps. A producer's accuracy of 81.6% for the cropland class is equivalent to an 18.4% error of omission. This means 18.4% of SE&NE Asia croplands were missing in the product. We report a lower user's accuracy (76.7%) for the cropland class for the continent, which translates to 23.3% commission error. By definition, this means that we over map 23.3% of the non-croplands as croplands. The decision to represent cropland extent with few omissions results in mapping some non-croplands as croplands.

Wider and more extended field visits to different parts of the continent could be helpful in improved training and validation for future map products. These and numerous other issues (e.g., implementing machine learning algorithms and uncertainties inherent in them) will continue to exist in cropland mapping over such expansive and diverse areas as SE and NE Asia. Nevertheless, advances made in this study are significant, especially in developing a nominal 30-m cropland extent of the entire SE and NE Asia with good accuracies.

Certain small islands in the Pacific are not classified and therefore data for these areas are not provided.

VII. Publications

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

1. Peer-reviewed publications relevant to this study

Oliphant, A., Thenkabail, P., Teluguntla, P., Xiong, J. Congalton, R., Yadav, K., 2017. Mapping cropland Extent of South East Asia using time-series Landsat 30-m data using Random Forest on Google Earth Engine (GEE) Cloud Computing. In Preparation.

Xiong, J., Thenkabail, P. S., James C. T., Gumma, M. K., Teluguntla, P., Congalton, R. G., Poehnelt, J., Kamini Yadav., et al. (2017). A Nominal 30-m Cropland Extent of Continental Africa Using Sentinel-2 data and Landsat-8 by Integrating Random Forest (SVM) and Hierarchical Segmentation Approach on Google Earth Engine. In press.

Xiong, J., Thenkabail, P. S., Gumma, M. K., Teluguntla, P., Poehnelt, J., Congalton, R. G., et al. (2017). Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS Journal of Photogrammetry and Remote Sensing, 126, 225–244.

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.

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,, 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. 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, 198: 490-503, 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.
, 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, 126: 225-244, https://doi.org/10.1016/j.isprsjprs.2017.01.019.

3. Web sites and Data portals:

https://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)
https://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. http://dx.doi.org/10.1016/j.jag.2008.11.002.

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. 65(1): 42-59. http://dx.doi.org/10.1016/j.isprsjprs.2009.08.004.

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 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) and The Asian Disaster Preparedness Center (ADPC). 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.

There were a number of International partners including The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

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 pthenkabail@usgs.gov

For the 30-m cropland extent product GFSAD30SEACE (Southeast & Northeast Asia), please contact:
Adam Oliphant aoliphant@usgs.gov
Prasad S. Thenkabail pthenkabail@usgs.gov
Pardhasaradhi Teluguntla pteluguntla@usgs.gov

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

X. Citations

Oliphant, A.J., Thenkabail, P.S., Teluguntla, P., Xiong, J., Congalton, R.G., Yadav, K., Massey, R., Gumma, M.K., 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 Southeast Asia 30 m V001* [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MEaSUREs/GFSAD/GFSAD30SEACE.001

XI. References

Bicheron, P., Amberg, V., Bourg, L., Petit, D., Huc, M., Miras, B., Brockmann, C., Hagolle, O., Delwart, S., Ranera, F., Leroy, M. and Arino, O., (2011). Geolocation assessment of MERIS GlobCover orthorectified products, IEEE Transactions on Geoscience and Remote Sensing. 49(8), 2972-2982. https://doi.org/10.1109/TGRS.2011.2122337

Bontemps, S., Defourny, P., Bogaert, E.V., Arino, O., Kalogirou, V. & Perez, J.R., (2010). GLOBCOVER 2009-Products description and validation report, UCLouvain and ESA Retrieved from http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf

Brovelli, M.A., Molinari, M.E., Hussein, E., Chen, J. & Li, R., (2015). The first comprehensive accuracy assessment of GlobeLand30 at a national level: Methodology and results. Remote Sensing. 7(4), 4191-4212. https://doi.org/10.3390/rs70404191

Bruinsma, J., (2011). The Resources Outlook: By How Much Do Land, Water and Crop Yields Need to Increase by 2050?, in *Looking Ahead in World Food and Agriculture: Perspectives to 2050*, ed. by Conforti, P. FAO, Rome. ISBN 978-92-5-106903-5

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, 7-27. https://doi.org/10.1016/j.isprsjprs.2014.09.002.

Cheng, Y., Yu, L., Cracknell, A.P. & Gong P., (2016). Oil palm mapping using Landsat and PALSAR: a case study in Malaysia. International Journal of Remote Sensing. 37(22), 5431-5442, https://doi.org/10.1080/01431161.2016.1241448

Congalton, R., (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment. 37, 35-46. https://doi.org/10.1016/0034-4257(91)90048-B.

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

Chan, R.H., Ho, C.W. & Nikolova, M., (2005). Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. IEEE Transactions on image processing. 14(10), 1479-1485. https://doi.org/10.1109/TIP.2005.852196.

Eliason, E.M. & McEwen, A.S. (1990). Adaptive box filters for removal of random noise from digital images. Photogrammetric Engineering and Remote Sensing. 56(4), 453-458.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L. and Seal, D., (2007). The shuttle radar topography mission. Reviews of geophysics. 45(2). https://doi.org/10.1029/2005RG000183.

Food & Agriculture Organization of the United Nations., (2013). FAO Statistical Yearbook 2013 World Food and Agriculture. Rome. ISBN 978-92-5-107396-4

Giri, C., Defourny, P., & Shrestha, S., (2003). Land cover characterization and mapping of continental Southeast Asia using multi-resolution satellite sensor data. International Journal of Remote Sensing. 24(21), 4181-4196, https://doi.org/10.1080/0143116031000139827

Google Earth Engine API, (2017) Landsat Algorithms, Retrieved from https://developers.google.com/earth-engine/landsat#simple-composite Last updated June 2, 2017.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R. (in press). Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, https://doi.org/10.1016/j.rse.2017.06.031

Gumma, M.K., Thenkabail, P.S., Teluguntla, P., Rao, M.N., Mohammed, I.A., & 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, 9, 981-1003.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R. & Kommareddy, A., (2013). High-resolution global maps of 21st-century forest cover change. Science. 342(6160), 850-853. https://doi.org/10.1126/science.1244693.

Hurni, K., Schneider, A., Heinimann, A., Nong, D.H., & Fox, J., (2017). Mapping the expansion of boom crops in mainland Southeast Asia using dense time stacks of Landsat data. Remote Sensing. 9(4), 320. https://doi.org/10.3390/rs9040320.

Kontgis, C., Schneider, A., & Ozdogan, M., (2015). Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. Remote Sensing of Environment. 169, 255-269. https://doi.org/10.1016/j.rse.2015.08.004.

Li, P., Feng, Z., & Xiao, C., (2017). Acquisition probability differences in cloud coverage of the available Landsat observations over mainland Southeast Asia from 1986 to 2015. International Journal of Digital Earth. 1–14. https://doi.org/10.1080/17538947.2017.1327619.

Oliphant, A.J., Wynne, R.H., Zipper, C.E., Ford, W.M., Donovan, P.F. & Li, J., (2017). Autumn olive (Elaeagnus umbellata) presence and proliferation on former surface coal mines in Eastern USA. Biological Invasions. 19(1), 179-195. https://doi.org/10.1007/s10530-016-1271-6.

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. https://doi.org/10.1016/j.rse.2016.10.010.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J.P., (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing. 67, 93-104. https://doi.org/10.1016/j.isprsjprs.2011.11.002.

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, JM., 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 Applied Earth Observation Geoinformatics. 38, 321-334. https://doi.org/10.1016/j.jag.2015.01.014.

Story, M. and R. Congalton. (1986). Accuracy assessment: A user's perspective. Photogrammetric Engineering and Remote Sensing. 52(3), 397-399.

Suepa, T., Qi, J., Lawawirojwong, S., Messina, JP., (2016). Understanding spatio-temporal variation of vegetation phenology and rainfall seasonality in the monsoon Southeast Asia. Environmental Research. 147, 621-629. https://doi.org/10.1016/j.envres.2016.02.005.

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, 211-261. doi:10.3390/rs2010211

Torbick, N., Chowdhury, D., Salas, W., & Qi, J., (2017). Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. Remote Sensing. 9(2), 119. https://doi.org/10.3390/rs9020119.

U.S Geological Survey. (2016). Landsat 8 (L8) Data Users Handbook, LSDS-1574, Version 2.0. Sioux Falls

U.S Geological Survey. (2017). Landsat 4-7 Surface Reflectance (LEDAPS) Product, Version 7.9. Sioux Falls

World Bank Group. (2016). Climate Change Knowledge Portal, Retrieved from http://sdwebx.worldbank.org/climateportal/index.cfm?page=global_map_region&ThisMap=AS