ISPRS Journal of Photogrammetry and Remote Sensing 144 (2018) 325–340

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform

Pardhasaradhi Teluguntla[a,b,*], Prasad S Thenkabail[a,*], Adam Oliphant[a], Jun Xiong[b], Murali Krishna Gumma[c], Russell G. Congalton[d], Kamini Yadav[d], Alfredo Huete[e]

[a] U.S. Geological Survey (USGS), 2255, N. Gemini Drive, Flagstaff, AZ 86001, USA
[b] Bay Area Environmental Research Institute (BAERI), NASA Research Park, Moffett Field, CA 94035, USA
[c] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad, India
[d] Department of Natural Resources and the Environment, University of New Hampshire, 56 College Road, Durham, NH 03824, USA
[e] University of Technology Sydney (UTS), PO Box 123, Broadway, NSW, Australia

ARTICLE INFO

Keywords:
Cropland mapping
Landsat
Machine learning algorithm
Random forest
Google Earth Engine
Australia
China

ABSTRACT

Mapping high resolution (30-m or better) cropland extent over very large areas such as continents or large countries or regions accurately, precisely, repeatedly, and rapidly is of great importance for addressing the global food and water security challenges. Such cropland extent products capture individual farm fields, small or large, and are crucial for developing accurate higher-level cropland products such as cropping intensities, crop types, crop watering methods (irrigated or rainfed), crop productivity, and crop water productivity. It also brings many challenges that include handling massively large data volumes, computing power, and collecting resource intensive reference training and validation data over complex geographic and political boundaries. Thereby, this study developed a precise and accurate Landsat 30-m derived cropland extent product for two very important, distinct, diverse, and large countries: Australia and China. The study used of eight bands (blue, green, red, NIR, SWIR1, SWIR2, TIR1, and NDVI) of Landsat-8 every 16-day Operational Land Imager (OLI) data for the years 2013–2015. The classification was performed by using a pixel-based supervised random forest (RF) machine learning algorithm (MLA) executed on the Google Earth Engine (GEE) cloud computing platform. Each band was time-composited over 4–6 time-periods over a year using median value for various agro-ecological zones (AEZs) of Australia and China. This resulted in a 32–48-layer mega-file data-cube (MFDC) for each of the AEZs. Reference training and validation data were gathered from: (a) field visits, (b) sub-meter to 5-m very high spatial resolution imagery (VHRI) data, and (c) ancillary sources such as from the National agriculture bureaus. Croplands *versus* non-croplands knowledge base for training the RF algorithm were derived from MFDC using 958 reference-training samples for Australia and 2130 reference-training samples for China. The resulting 30-m cropland extent product was assessed for accuracies using independent validation samples: 900 for Australia and 1972 for China. The 30-m cropland extent product of Australia showed an overall accuracy of 97.6% with a producer's accuracy of 98.8% (errors of omissions = 1.2%), and user's accuracy of 79% (errors of commissions = 21%) for the cropland class. For China, overall accuracies were 94% with a producer's accuracy of 80% (errors of omissions = 20%), and user's accuracy of 84.2% (errors of commissions = 15.8%) for cropland class. Total cropland areas of Australia were estimated as 35.1 million hectares and 165.2 million hectares for China. These estimates were higher by 8.6% for Australia and 3.9% for China when compared with the traditionally derived national statistics. The cropland extent product further demonstrated the ability to estimate sub-national cropland areas accurately by providing an R^2 value of 0.85 when compared with province-wise cropland areas of China. The study provides a paradigm-shift on how cropland maps are produced using multi-date remote sensing. These products can be browsed at www.croplands.org and made available for download at NASA's Land Processes Distributed Active Archive Center (LP DAAC) https://www.lpdaac.usgs.gov/node/1282.

* Corresponding authors at: U.S. Geological Survey (USGS), 2255, N. Gemini Drive, Flagstaff, AZ 86001, USA (P. Teluguntla).
E-mail addresses: pteluguntla@usgs.gov (P. Teluguntla), pthenkabail@usgs.gov (P.S. Thenkabail).

https://doi.org/10.1016/j.isprsjprs.2018.07.017
Received 12 February 2018; Received in revised form 26 July 2018; Accepted 27 July 2018
Available online 10 August 2018

1. Introduction

Accurate, and precise agricultural cropland extent products over very large areas that map small to large farms are of great importance to assess and monitor global food and water security. They are a critical part of land system studies (Verburg et al., 2013). Such products are also of great importance for assessing global crop water use, crop productivity (productivity per unit of land), water productivity (productivity per unit of water or crop per drop), and food security studies (Foley et al., 2011; Thenkabail et al., 2010; Teluguntla et al, 2015b; Matejicek and Kopackova, 2010). Remote sensing based spatially distributed cropland maps with high spatial resolution provide an efficient way to monitor croplands (Foley et al., 2011; Fritz et al., 2015; Yu et al., 2013). Over the last two decades, several global and regional cropland products have been produced using medium to coarse resolution (250-m to 1-km) remote sensing data such as the Advanced Very High-Resolution Radiometer (AVHRR) and the Moderate-Resolution Imaging Spectroradiometer (MODIS) data (Biradar et al., 2009; Kumar et al., 2018; Thenkabail et al., 2009, 2012; Pittman et al., 2010; Portmann et al., 2010; Siebert and Döll, 2010; Salmon et al., 2015; Waldner et al., 2015, 2016). These products are very useful for a preliminary understanding of agricultural croplands in terms of their spatial distribution patterns and their characteristics such as crop dominance and cropping intensities. However, the coarse resolution of these products limits their usefulness in assessing small agriculture fields (Teluguntla et al., 2015b; Thenkabail et al., 2010). Further, there are several global to regional land use/land cover (LULC) products produced using multiple remote sensing data in which agricultural croplands is one or more classes. Some examples are: DIScover (Loveland et al., 2000); GLC500m (Friedl et al., 2010); MCD12Q1 (Liang et al., 2015); Globecover (Defourny et al., 2009); FROM-GC (Gong et al., 2013); FROM-GLC (Yu et al., 2013); and Globeland30 (Arsanjani et al., 2016; Chen et al., 2015). However, these products were focused on LULC in which mapping croplands in detail was not the primary objective. Hence, the cropland accuracies suffer (Yang et al., 2017). Further, most of these products are also coarse resolution. Most of these products fail to map individual farm fields, especially when they are small and\or fragmented. Definitions of croplands also vary from product to product, resulting in different results of cropland extent and their characteristics in each of these products. Overall, existing cropland extent products are coarse resolution, lack field level details, and/or are mapped as part of other LULC classes where specific cropland class focus is missing. As a result, uncertainties and errors in cropland locations are very high.

In the past, number of advanced remote sensing methods have been used for mapping agricultural croplands. These studies were conducted using data from multiple sensors across many spatial, spectral, radiometric, and temporal resolutions for both irrigated and rainfed crops (Biggs et al., 2006; Dheeravath et al., 2010; Funk and Brown, 2009; Gumma et al., 2011. 2016; Ozdogan and Woodcock, 2006; Pervez et al., 2014; Teluguntla et al. 2015a; Thenkabail et al., 2009, 2012; Velpuri et al., 2009; Xiao et al., 2006). These methods consist of pixel-based, object-based, or a combination of both approaches that used either supervised or unsupervised classification techniques. Pixel-based approaches include: a) Random forest algorithm (Tatsumi et al., 2015; Wang et al., 2015; Gislason et al, 2006); (b) Support vector machines (Mountrakis et al., 2011; Shao and Lunetta, 2012); (c) decision tree algorithms (Ozdogan and Gutman, 2008; Waldner et al., 2016); (d) Tassel cap brightness-greenness-wetness (Cohen and Goward, 2004; Gutman et al., 2008; Masek et al., 2006); (e) Spectral matching techniques (Gumma et al., 2014; Thenkabail et al., 2007; Teluguntla et al., 2017a); (f) Phenological approaches (Pan et al., 2015; Teluguntla et al. 2015a; Zhong et al., 2016; Zhou et al., 2016); (g) the Automated Cropland Classification Algorithms (Thenkabail and Wu, 2012; Teluguntla et al., 2017a; Waldner et al., 2015); and (h) Machine learning programming involving a combination of multiple methods (DeFries and Chan, 2000; Duro et al., 2012; Pantazi et al., 2016). Object-based approaches (Peña-Barragán et al., 2011; Peña et al., 2014) include Hierarchical Image Segmentation software or HSeG (Tilton et al., 2012). A combination of pixel-based and object-based methods have also been recently attempted (Xiong et al., 2017a; Chen et al., 2018). However, these methods and approaches were overwhelmingly applied on: (a) multi-temporal moderate resolution (250-m or higher) remotely sensed data, and/or (b) small areas, and/or (c) high-resolution (Landsat 30-m) remotely sensed data with limited multi-temporal images.

Hitherto, availability of cloud-free, high quality images as well as use of multi-temporal, high-resolution data over very large areas for cropland mapping has been challenging and resource intensive. However, these challenges have been overcome through a paradigm shift in remote sensing data collection, management, and processing. First, Landsat-8 Operational Land Imager (OLI) data and Landsat-7 Enhanced Thematic Mapper+ (ETM+) data at 30-m spatial resolution were utilized every 16-days for 3-years (2013–2015) for entire Australia and China. Managing massive volumes of Landsat data for analysis over very large areas is a big challenge when adopting traditional remote sensing approaches that use commercial imaging processing software on workstation PC based systems. No matter how powerful the systems are, the entire process of data analysis including, pre-processing, over very large areas involving 1000's of Landsat images is cumbersome, slow, and tedious. However, in the current era of adopting powerful machine learning algorithms (MLAs) in cloud computing environments such as Google Earth Engine (GEE) these limitations have been overcome allowing planetary scale remote sensing at high spatial resolutions as illustrated by Erickson (2014) and Gorelick et al. (2017). Gorelick et al. (2017) demonstrated that multi-petabyte archive of georeferenced datasets can be combined in the GEE catalog that includes images from Earth observing satellites and airborne sensors (e.g., USGS Landsat, NASA MODIS, USDA NASS CDL), weather and climate datasets, as well as digital elevation models. This system has exceptional data organization, and has enabled geo spatial processing over very large areas. Along with computing and storage resources, GEE also supports major MLAs useful for image enhancement and, image classification, and allows batch processing through JavaScript or Python on Application Program Interfaces (APIs). These capabilities reduce most of preprocessing steps needed in traditional remote sensing approaches. Very recently, several studies have used the GEE platform for large-scale (continental, global) mapping (Xiong et al., 2017a, 2017b).

Thereby, the overarching goal of this study was to map cropland extent in detail (e.g., showing all individual farms whether small or big) using high-resolution (30-m) multi-year (2013–2015) time-series (16-day) Landsat-8 OLI data over the entirety of China and Australia. These two countries have very large cropland areas with distinct cropping systems. Australia is a major agriculture producer and exporter with large scale industrial farm fields. Pastoral farming is another major agricultural land use in Australia. Australian farmers produce cereals, legumes and oilseeds on a large scale (ABARES, 2016) for human consumption and livestock feed. In contrast, an overwhelming proportion of Chinese farms are small, fragmented, but often contiguous over large areas due to intensive land use for agriculture. Average crop field size in China is less than a hectare (Samberg et al., 2016); fine resolution satellite data is required to map such crop fields. Landsat data with 30-m spatial resolution is ideal dataset to map cropland extent in China. Whereas the average crop field size in Australia is 100 ha (Samberg et al., 2016) which is much larger than most crop fields in China. Medium resolution sensor such as MODIS at 250-m (1 pixel is approximately 6.25 ha) are highly inadequate to map smaller and/or fragmented crop fields. High-resolution 30-m (1 pixel is approximately 0.09 ha) Landsat-8 OLI 30-m imagery is expected to map small and fragmented farms in addition to large farms. Chinese farmlands are also very diverse, spread across mountains, river banks, and large plains. China's agriculture feeds 1.38 billion people whereas Australia's much smaller population allows it to be a major exporter of food. China is the

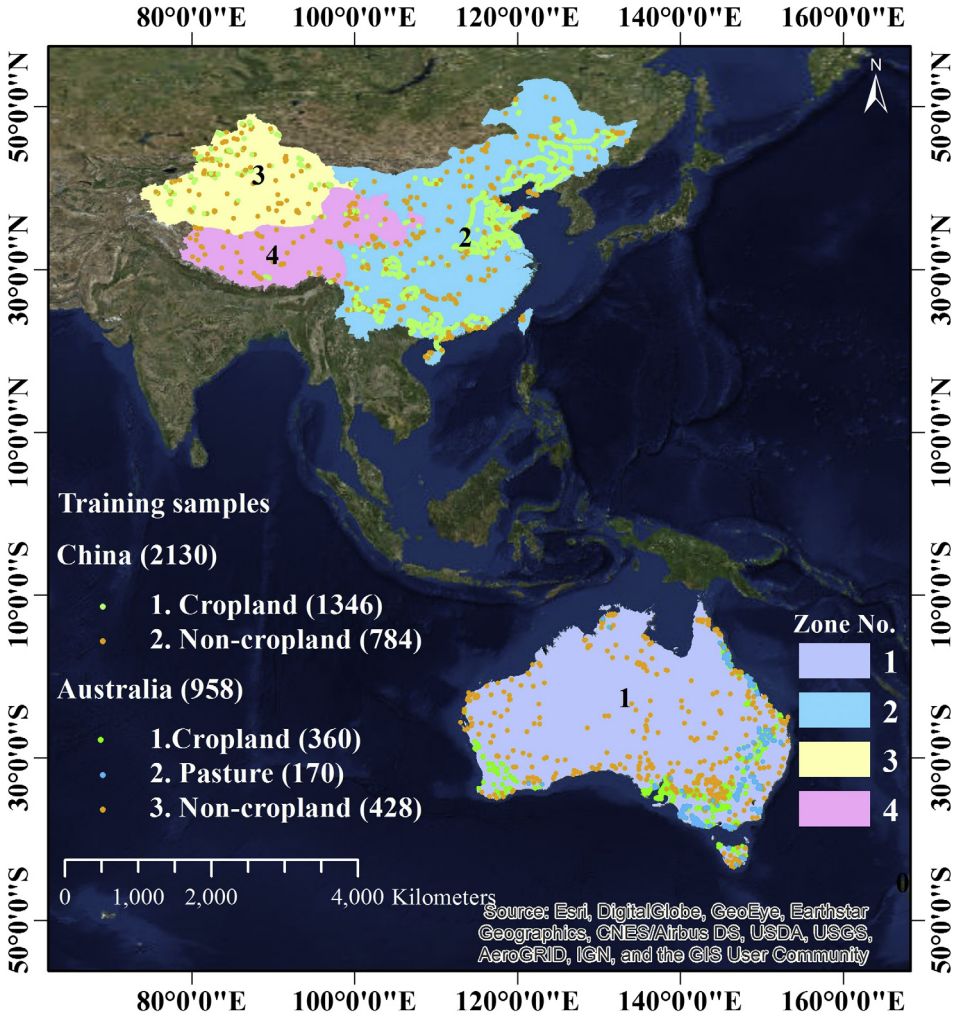

Fig. 1. Stratification of the study area into distinct and broad refined agro-ecological zones (RAEZs). The figure also shows the distribution of the reference training data used in the Random Forest machine-learning algorithm. The Random Forest pixel-based supervised machine learning algorithm used in this study was trained using reference training data falling within each of these zones to separate croplands from non-croplands.

topmost producer of three major food crops (rice, wheat, and maize) in the world (FAOSTAT). The best cropland products of Australia and China known by the authors are derived from other LULC products. For example, for Australia, croplands were mapped as part of land cover products such as Dynamic Land Cover Data (DLCD) developed by Geo Science Australia (Lymburner et al., 2011). Another cropland product for Australia is from the Australian Collaborative Land Use and Management Program (ACLUMP) by the Department of Agriculture and Water resources (http://www.agriculture.gov.au/abares/aclump). However, there is no unique cropland layer available for Australia and China like United States Department of Agriculture (USDA) National Agricultural Statistical Service (NASS) 30 m Cropland Data Layer (CDL) derived from Landsat for USA (Johnson and Mueller, 2010; Boryan et al., 2011; Han et al., 2012). Due to the above reasons, mapping croplands at high-resolution of these two countries is of great importance.

Thereby, specific objectives of the study were to develop a precise 30-m cropland extent map of Australia and China using 16-day Landsat-8 OLI data for the nominal year 2015 using an RF classifier through the GEE cloud-computing platform. We chose a pixel-based RF MLA approach for this study because MLAs have been particularly useful in classifying large datasets with high spatial and temporal resolutions for large area land cover mapping (Pelletier et al., 2016). Second, this study used a large collection of reference training and validation data obtained from: (a) ground-data through extensive field visits, (b) samples interpreted from sub-meter to 5-m very high-resolution imagery data, and (c) data sourced from reliable secondary sources. These reference datasets were invaluable in training the RF classifier, as well as to assess classification accuracies, and establish uncertainties. Third, computed cropland area statistics from the 30-m cropland product were compared with national and sub-national cropland area statistics obtained from sources such as the Food and Agricultural Organization (FAO) of the United Nations (UN) and National Governmental sources.

2. Study area

Our study area consisted of two major cropland areas of the world: China and Australia. The two regions were further stratified into four refined agro-ecological zones (RAEZs) (Fig. 1), which helped identify areas of similar farming practices, soil types, and climatic patterns. These RAEZs were modifications of the Food and Agricultural Organization's (FAO's) global agro-ecological zones (GAEZs) which are based on the number of growing degree days, soil, and terrain data, all provided at 10-km spatial resolution (Fischer et al., 2000). However, GAEZs has too many zones for cropland classification purposes as many of these zones only have a very small proportion of crops relative of the overall land area of the zone. Therefore, we refined GAEZs into RAEZs by utilizing Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 2 (GDEM V2) 30-m data (https://www.asterweb.jpl.nasa.gov/gdem.asp), slope

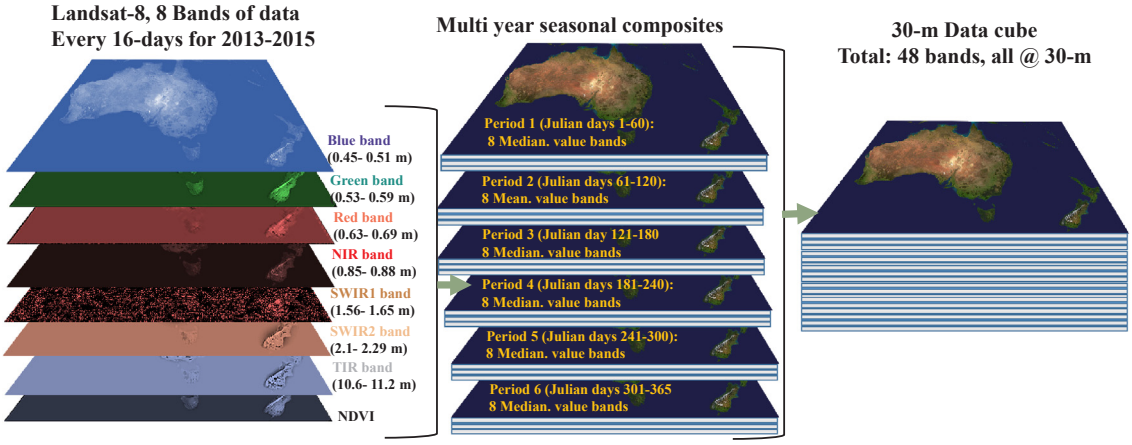

Fig. 2a. 30-m Data-cube for Australia composited for 6 time-periods using 2013–2015 Landsat-8 data. For each period (e.g., period 1: Julian days 1–60), eight bands (blue, green, red, NIR, SWIR1, SWIR2, TIR1, and NDVI) were composited, taking median value of a given pixel over the period. From 6 periods, there was a 48-band mega-file data cube in Google Earth Engine collection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

derived off 30-m GDEM data, and the proportion of croplands in a zone. A number of smaller RAEZs were combined into larger zones depending on the zone's importance for croplands. This resulted in three broad RAEZs across China, and a single RAEZ for Australia (Fig. 1). Classification using MLAs was carried out in each of these RAEZs separately and then similar classes across zones were combined.

3. Datasets

Two major types of data were used in the study. First, we will describe and discuss the satellite sensor data followed by reference training and validation data of Australia and China.

3.1. Satellite imagery: Landsat-8 and Landsat-7

In order to cover crop dynamics in different time-periods, Landsat-8 OLI Top of Atmosphere (TOA) 16-day satellite sensor data collections available on GEE cloud for three years (2013–2015) were organized for Australia and China. In addition, Landsat-7 was used to fill some data gaps (Irons et al., 2012) over 2013–2015 for China, aiming to provide seamless cloud-free 30-m wall-to-wall data coverage (Fig. 2a). Specific image collections used were LANDSAT/LC8_L1T_TOA and LANDSAT/LE7_L1T_TOA for 2013–2015 (sometimes referred to as nominal 2015). The Landsat TOA data available on GEE are orthorectified and corrected for solar angle (USGS TOA product guide). The coefficients were derived from USGS TOA product guide and Chander et al. (2009). There is a 16-day revisit time for Landsat-8 OLI and Landsat-7 ETM+ 30-m data. It is not possible to get continuous 16-day cloud free time-series data for wall-to-wall coverage for any part of the region due to cloud cover. 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., Figs. 2a and 2b). Finally, 30-m mega-file data-cubes (MFDCs) were created as per the following steps leading to a 48-band MFDC (Fig. 2a) of Australia for 6 time-periods (time composites) and a 32-band MFDC of China (Fig. 2b) for 4 time-periods.

We used the multi-year (2013–2015) 16-day Landsat-8 OLI (Fig. 2a) and Landsat-7 ETM+ (Fig. 2b) data for the study area to: (1) ensure wall-to-wall coverage of data, and (2) reduce the effects of cloud cover. The nominal year 2015 was further divided into multiple intervals or periods (Figs. 2a and 2b) depending on the seasonal variability in the region and availability of the cloud free Landsat data (Table 1). Bi-monthly or tri-monthly MFDC cloud-free wall-to-wall collections (Figs. 2a and 2b) of Australia and China were established.

For Australia, cloud-free or near cloud-free images could be created for bi-monthly time-periods (Fig. 2a) leading to a total of 6 time-periods (period 1: Julian days 1–60, period 2: 61–120, period 3: 121–180, period 4: 181–240, period 5: 241–300, period 6: 301–365; Fig. 2a) over

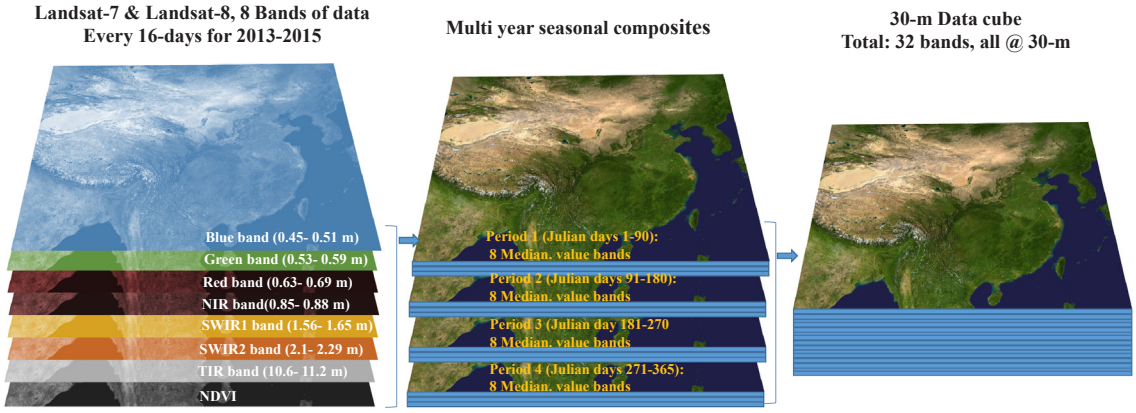

Fig. 2b. 30-m Data-cube for China composited for 4 time-periods using 16-day data of Landsat-8 and Landsat-7 for the years 2013–2015 (e.g., period 1: Julian days 1 to 90). For each period eight bands (blue, green, red, NIR, SWIR1, SWIR2, TIR1 and NDVI) were composited, taking the median value of a given pixel over the period. From the 4 time-periods, there was a 32-band mega-file data cube in the Google Earth Engine collection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Characteristics of multi-temporal multi-year Landsat 30-m data used in the study.

Country name	Landsat image series	Years of data	Julian days over data are time-composited[a]	Bands used[b] for each composited period	Total #of bands used in mega-file data cube	Data provider name
Australia	Landsat-8	2013, 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, TIR1, and NDVI	48	USGS
China	Landsat-7 and 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, TIR1, and NDVI	32	USGS

Then each band (e.g., blue) is derived using the median value from these 4 images. Similarly, for all bands.
NDVI = normalized difference vegetation index
[a] 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.
[b] NIR = near-infrared, SWIR = short-wave infrared, TIR = thermal infrared.

a 12-month period. In comparison, tri-monthly time-periods were created for China (Fig. 2b) because a greater number of days were required 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 only 4 time-periods (period 1: Julian days 1–90, period 2: 91–180, period 3: 181–270, period 4: 271–365; Fig. 2b) over 12-month period. It should be noted that for each time-period (e.g., 1–60 days) Landsat data from multiple years (2013–2015) were used to maximize the chances of pure cloud free pixels over Australia and China. As a result, the process involved gathering all the Landsat-8 16-day images over Australia (Fig. 2a), and all the Landsat-8 as well as Landsat-7 images over China (Fig. 2b) available for each time-period (e.g., period 1: 1–60 Julian days). The following bands and indices were used for each time-period: blue (0.45–0.51 μm), green (0.53–0.59 μm), red (0.63–0.69 μm), NIR (0.85–0.89 μm), SWIR1 (1.55 1.65 μm), SWIR2 (2.1–2.3 μm), and TIR1 (10.60–11.19 μm); (Table 1, Figs. 2a and 2b) along with Normalized Difference Vegetation Index (NDVI) were time-composited by taking median value of each pixel of each band (Roy et al., 2014). Thereby, for Australia, eight median value bands composed over 6 time-periods resulted in a 48-band MFDC (Fig. 2a). Whereas for China, eight median value bands composed over 4 time-periods resulted in a 32-band MFDC (Fig. 2b). The band stack and time-periods leading to MFDC are shown in Table 1 as well as in Figs. 2a and 2b. All compositions were performed on the GEE cloud-based geospatial platform for planetary-scale data analysis (Gorelick et al., 2017). Landsat TOA products were used instead of surface reflectance (SR) due to the limited temporal availability of Landsat-7 and Landsat-8 SR imagery on GEE for our study regions.

3.2. Reference croplands samples for training and validation

Reference training and validation data were collected for Australia

Fig. 3a. Ground data for Australia. Illustration of some of the sample reference training and validation data collected during field visit in Australia.

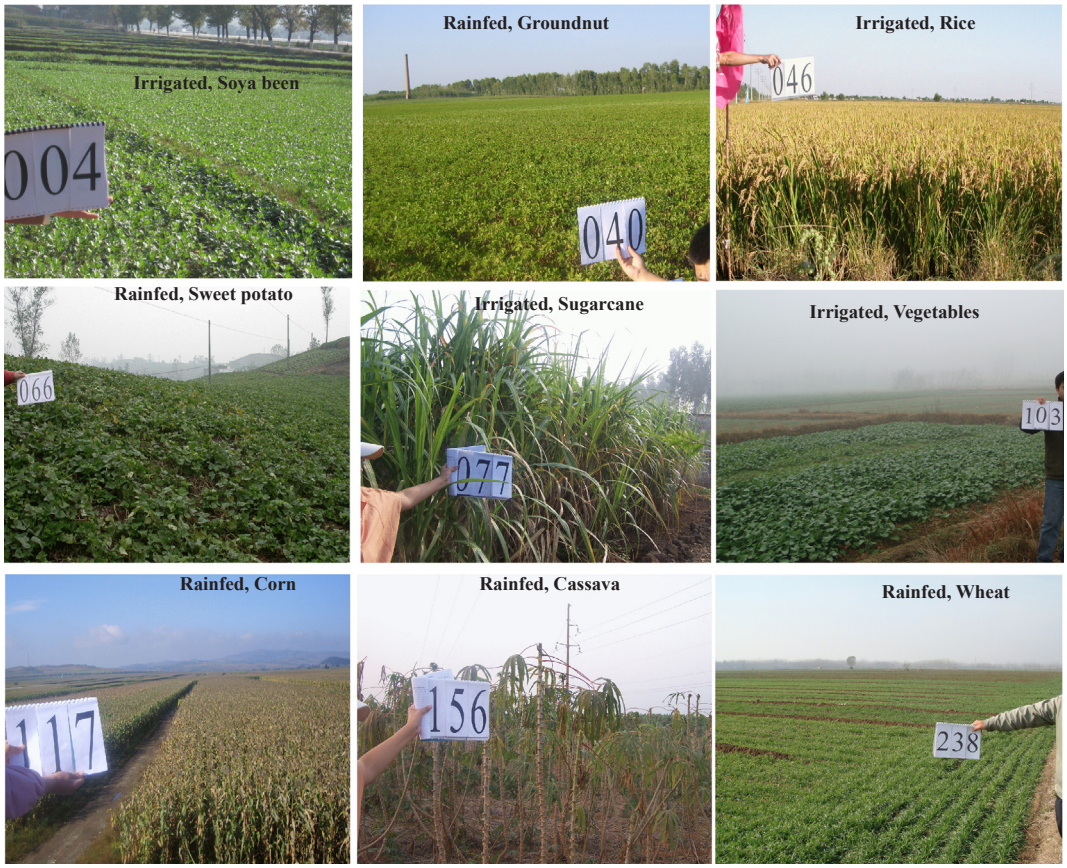

Fig. 3b. Ground data for China. Illustration of some of the sample reference training and validation data collected during field visit in China obtained in collaboration with Chinese Academy of Agricultural Sciences (CAAS).

and China, from three high quality sources as described below. Part of the training data was used to gather knowledge and train the MLA and the other part was used to test the classes. Validation data were used to assess accuracies, errors, and uncertainties. Data were collected keeping the cropland definition in mind. Croplands were defined in this study as agricultural lands with annual standing crops + cropland fallows + permanent crops (e.g., coffee, tea, rubber) (Teluguntla et al., 2015b). Samples were collected to cover wide range of cropland and non-cropland classes (e.g., Figs. 3a and 3b) with excellent spatial distribution in Australia (Fig. 4a) and China (Fig. 4b). Since the satellite data were collected for the 2013–2015 time-period there are possibilities that croplands in one year or season may remain cropland fallows in another year or season (Gumma et al., 2018). The cropland areas of Australia and China are well established. Hence croplands changing to non-croplands in Australia and China in such short time period (2013–2015) was minimal and was estimated by us to be < 1% of the total cropland areas of these countries. Further, we also observed the multi-year NDVI patters of the sample areas to see whether the sample locations were “pure” (unchanged) or “changed” (showed significant change over time). Reference training and validation data were then appropriately labeled.

3.2.1. Ground reference and validation data through field visits

First, we conducted an extensive field survey during September and October of 2014, the peak crop-growing season for crops in Australia. More than 4400 ground samples (e.g., Figs. 3a and 4a) were collected from New South Wales (NSW), Victoria (VIC), South Australia (SA), and Western Australia (WA) states 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), some cropland fallow samples and non-cropland samples. Such data collection over a wide range of crop types helped us to capture variability across croplands and helped in the development of robust training datasets. A sample illustration of the ground reference data acquired across Australia is shown in Fig. 3a and the spatial distribution of data across Australia is shown in Fig. 4a. The samples shown in Fig. 4a cover the entire agricultural landscape of Australia except the northern part of the country where we used sub-meter to 5-m data.

Similarly, we obtained ground reference data for China through collaboration with the Chinese Academy of Agricultural Sciences (CAAS) between 2008 and 2014 time-period. These field data were spatially well distributed throughout agricultural cropland areas of China. A total of 2120 ground samples (e.g., Figs. 3b and 4b) 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. A sample illustration of the ground reference data distribution across China is shown in Fig. 3b. For both Australia and China, there was a minimum of one-kilometer distance between each sample to ensure spatial autocorrelation do not occur (Congalton, 2015). We further verified all ground data samples for croplands *versus* non-croplands using sub-meter to 5-m very high-resolution data (also see Section 3.2.2).

3.2.2. Very high spatial resolution imagery (VHRI) data

Second, some samples were obtained throughout Australia, and China by interpreting sub-meter to 5-meter very high spatial resolution imagery (VHRI) data of 4 recent years (2012–2015) available to US

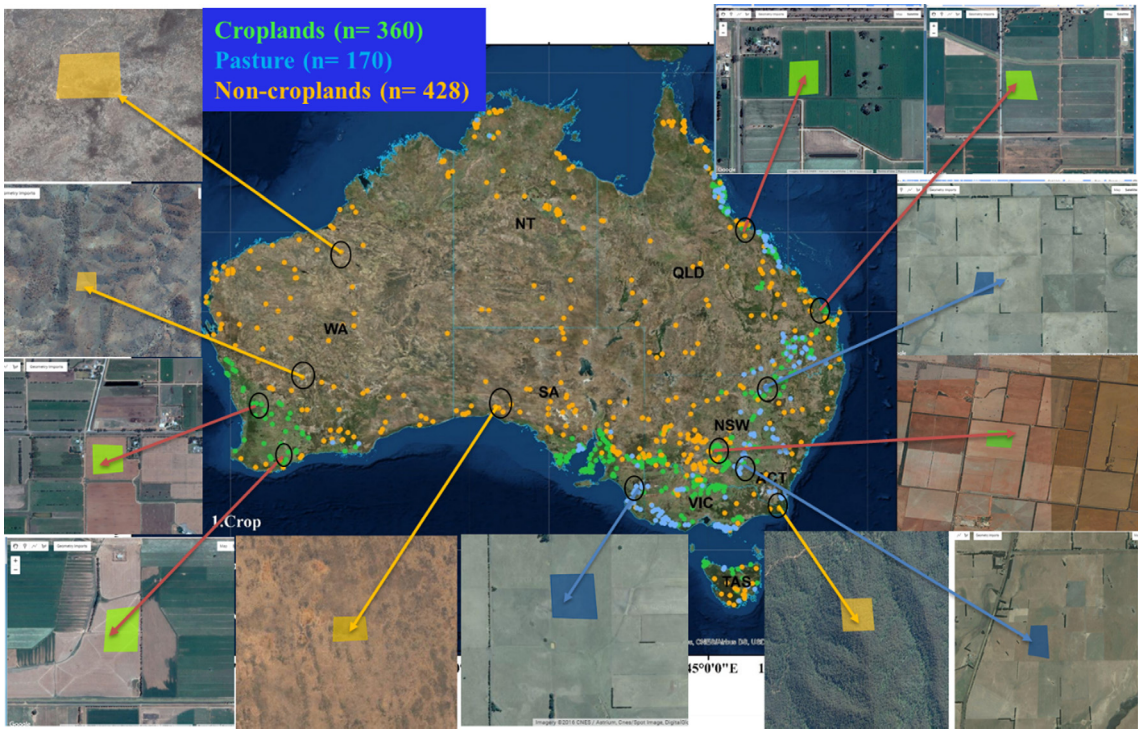

Fig. 4a. Sub-meter to 5-m very high-resolution imagery data for Australia. Illustration of the reference training and validation data of Australia collected using sub-meter to 5-m very high spatial resolution imagery.

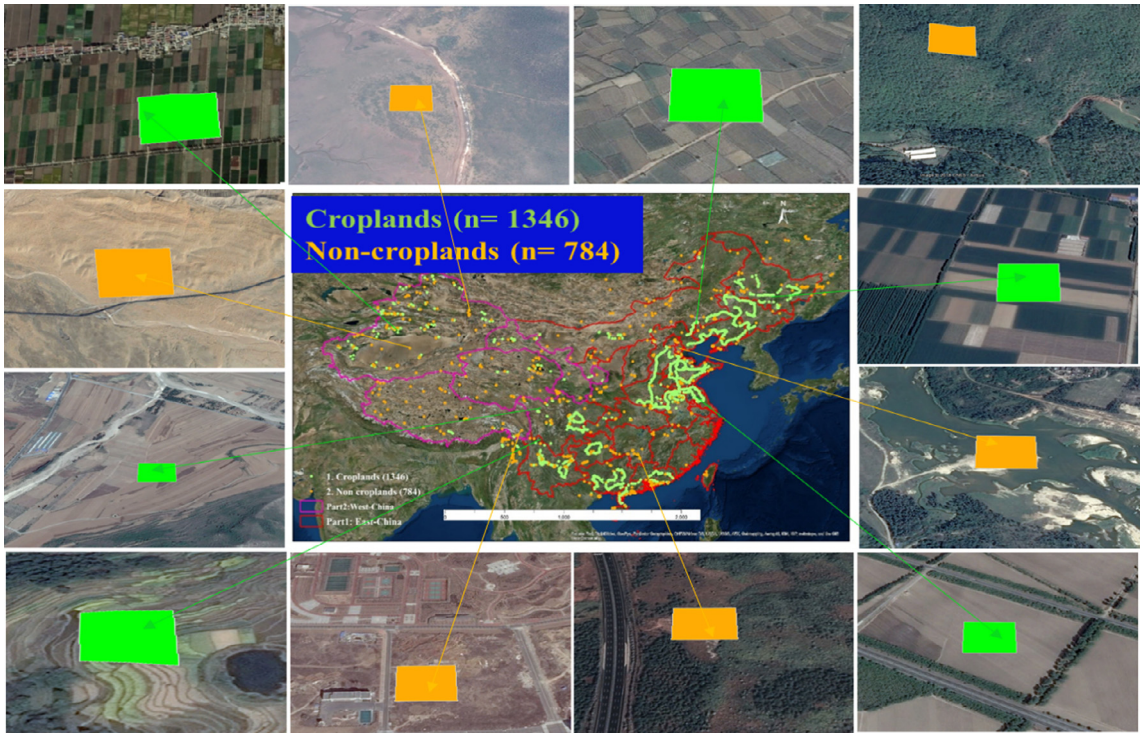

Fig. 4b. Sub-meter to 5-m very high-resolution imagery data for China. Illustration of the reference training and validation data of China collected using sub-meter to 5-m very high spatial resolution imagery.

Government entities through the National Geospatial Agency (NGA) (Marshall and Thenkabail, 2015) and multiple years of imageries obtained from Google Earth Pro. From this, we collected 1210 reference and 1710 validation samples (e.g., Fig. 4a for Australia, Fig. 4b for China).

3.2.3. Tertiary sources

Third, reference data were obtained from national systems and other reliable sources such as Geo Science Australia, Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES).

A summary of sample distribution of the reference training and validation data for Australia and China is provided in Table 2. The

Table 2
Reference training and validation data. A number of reference samples used for training the Random Forest machine-learning algorithm and number of validation samples used for independent accuracy assessment.

Country	Class	Training samples	Validation samples
Australia	Cropland	360	80
	Pasture	170	
	Non-cropland	428	820
	Total	958	900
China	Cropland	1346	340
	Non-cropland	784	1632
	Total	2130	1972

reference data were divided into two distinct datasets: roughly 50% for training and 50% for validation (Table 2). The training data were used to develop knowledge base and train the machine-learning algorithms (e.g., Random Forest in this study). The validation data were used to perform accuracy assessments leading to errors of omissions, errors of commissions, and overall accuracies.

4. Methodology

The goal of this study was to produce accurate Landsat 30-m derived cropland extent product for Australia and China. We used a pixel-based supervised classification approach with the RF classifier on the GEE cloud computing platform to develop cropland extent product for Australia, and China using Landsat 30-m, 16-day time-series data (Figs. 2a and 2b) for the 2013–2015 time-period. An overview of the methodology is shown in Fig. 5.

4.1. Random forest machine learning algorithm

We selected the pixel-based supervised Random Forest (RF) machine learning algorithm because it is generally immune to data noise and overfitting and is extremely useful in classifying remote sensing data. Furthermore, RF classifiers can successfully handle high data dimensionality and typically achieve higher accuracies in comparison with other approaches such as maximum likelihood, single decision trees, and single layer neural networks (Belgiu and Drăguţ, 2016; Lawrence et al., 2006; Na et al., 2010). Random Forest algorithm is a machine learning method, in which the RF classifiers construct multiple de-correlated random decision trees that are bootstrapped and aggregated to classify a dataset by using the mode of predictions from all decision trees (Breiman, 2001). The RF classifier is more robust than

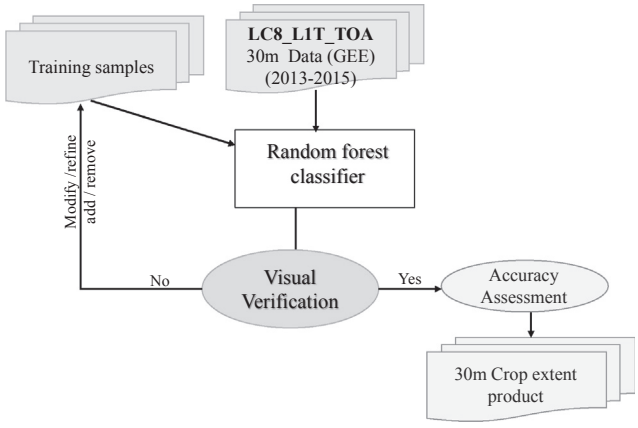

Fig. 5. Overview of methodology for cropland mapping. This study used a pixel-based supervised Random Forest machine learning algorithm for classification. Analysis executed on Google Earth Engine cloud-computing platform.

single decision tree (Chan and Paelinckx, 2008) and easier to be implemented than many other advanced classifiers such as Support Vector Mission (SVM) (Pelletier et al., 2016) as recently established by Xiong et al. (2017a). It 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. Random Forest classifiers also provide a quantitative measurement of each variable's contribution to the classification output, which is useful in evaluating the importance of each variable. They provide an internal accuracy assessment by using an 'out-of-bag' (OOB) technique, in which about a third of the data is kept aside as a validation dataset to assess accuracy of the classification. This technique can be used to cross-validate the RF classifier using independent datasets.

Random Forest classifiers available in GEE use six input parameters: (1) number of classification trees, (2) number of variables used in each classification tree, (3) minimum leaf population, (4) bagged fraction of the input variables per decision tree, (5) out-of-bag mode, and (6) random seed variable for decision tree construction. When the number of trees increases, the overall accuracy of classification increases without overfitting (Breiman, 2001). While training sample imbalance can affect the RF classification output by over-fitting the majority class (Breiman, 2001; Chen et al., 2004), various methods such as down-sampling the majority class or duplicating the minority samples can provide immunity against over-fitting (Sun et al., 2007). Further, the choice of optimized parameter values using the OOB outputs can eliminate overfitting in the RF classifier. The optimized parameter values were selected by selecting the training samples (Table 2), running the RF algorithm, and testing the classification output for overall, producer's and user's accuracies in error matrix. The goal is to obtain not just the high overall accuracies, but also a good balance of producer's accuracies (or least errors of omissions) and user's accuracies (or least errors of commissions). It is not just the high number of training samples of a class that help attain optimal accuracies, but the purity (e.g., pure cropland samples instead of mixed) of the samples as well.

All supervised pixel-based classifications rely heavily on the input training samples. To discriminate croplands under various environments and conditions, two criteria are very important:

1. RF classifications need to take place in AEZs or RAEZs (e.g., Fig. 1),
2. The sample size (Table 2) of the initial training dataset for the RF classifier needs to be large, especially in complex regions.

All samples were selected to represent a 90-m × 90-m polygon (Table 2 and Section 3 and its sub-sections). 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 2). Second, we utilized the ground data from previous efforts for China and other reliable sources. Third, sub-meter to 5-m VHRI, available to us for the entire study region, was used to generate croplands *versus* non-cropland interpretations by multiple analyses across China, Australia, 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 RF machine-learning algorithm as illustrated in Fig. 5 and described step-by-step below:

1. Build RF classifier using existing training samples.
2. Based on the established classifier, classify 30-m MFDC (Figs. 2a and 2b) using the RF algorithm on GEE cloud;
3. Visual assessment of classification results is compared with existing reference maps as well as sub-meter to 5-m VHRI;
4. Added 'crop' samples in missing area and 'non-crop' samples to over classified areas by referencing sub-meter to 5-m VHRI 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 × 90-m polygon.

5. Loop steps 1–4 with enlarged training dataset until classification becomes stable and high classification accuracy is achieved.

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 (Fig. 1) to carry out classification (Fig. 5): the iterative selection was repeated ~4–5 times, in each of the 3 zones, to improve the initial classified results. Initially, we started with a small number of samples (3 0 0) and slowly increased the sample size till we reached a high level of accuracies. After each iteration, we visually compared the classification result with sub-meter to 5-m VHRI, randomly, at 100 s of locations. If the classification results were not satisfactory, we increased the training samples till we attained satisfactory classification results. Once that was achieved, the accuracy assessment team performed the accuracies using the independent validation data (Table 2). A product was accepted as final only when the overall, producer's and user's accuracies were adequately high (typically above 80%).

4.2. Croplands versus non croplands knowledge capture

Knowledge is very important to train a RF algorithm. First knowledge (e.g., Figs. 6a and 6b) was captured (i.e., extracted values from all bands used in this study) using training-samples (Table 2, Sections 3.2.1–3.2.3) from Landsat data cube (Figs. 2a and 2b). For illustration, we have shown knowledge distinguishing croplands, non-croplands, and pastures for all training samples for Australia plotted in a box plot (e.g., Fig. 6a). This shows which bands and periods are able to separate croplands from non-croplands. For example, Fig. 6a shows which period is more appropriate to separate croplands from non-croplands. The "Zoom-in" part of NDVI clearly shows that separability of croplands from non-croplands is maximum during period 4 and period 5, which are during the peak growing season of crops in Australia. Similarly, separability of croplands from pastures is maximum during period 3 and period 6.

4.3. Cloud computing on Google Earth Engine/processing platforms

We used GEE cloud computing for the pixel-based RF machine learning algorithm for cropland classification. The Landsat archive available on GEE is already pre-processed for atmospheric and topographic effects, which saved us much effort in data download and pre-processing time. We used JavaScript API in the GEE code editor. The zonal boundaries and all training samples were imported into GEE via Google fusion tables.

5. Results and discussions

There are four distinct components under this section. It starts with a presentation of the 30-m Landsat derived cropland extent product for Australia and China. This is followed by an assessment of accuracies, errors, and uncertainties of the cropland extent products. Then cropland areas, at the national and sub-national level are presented and discussed. Finally, the cropland extent product is spatially compared

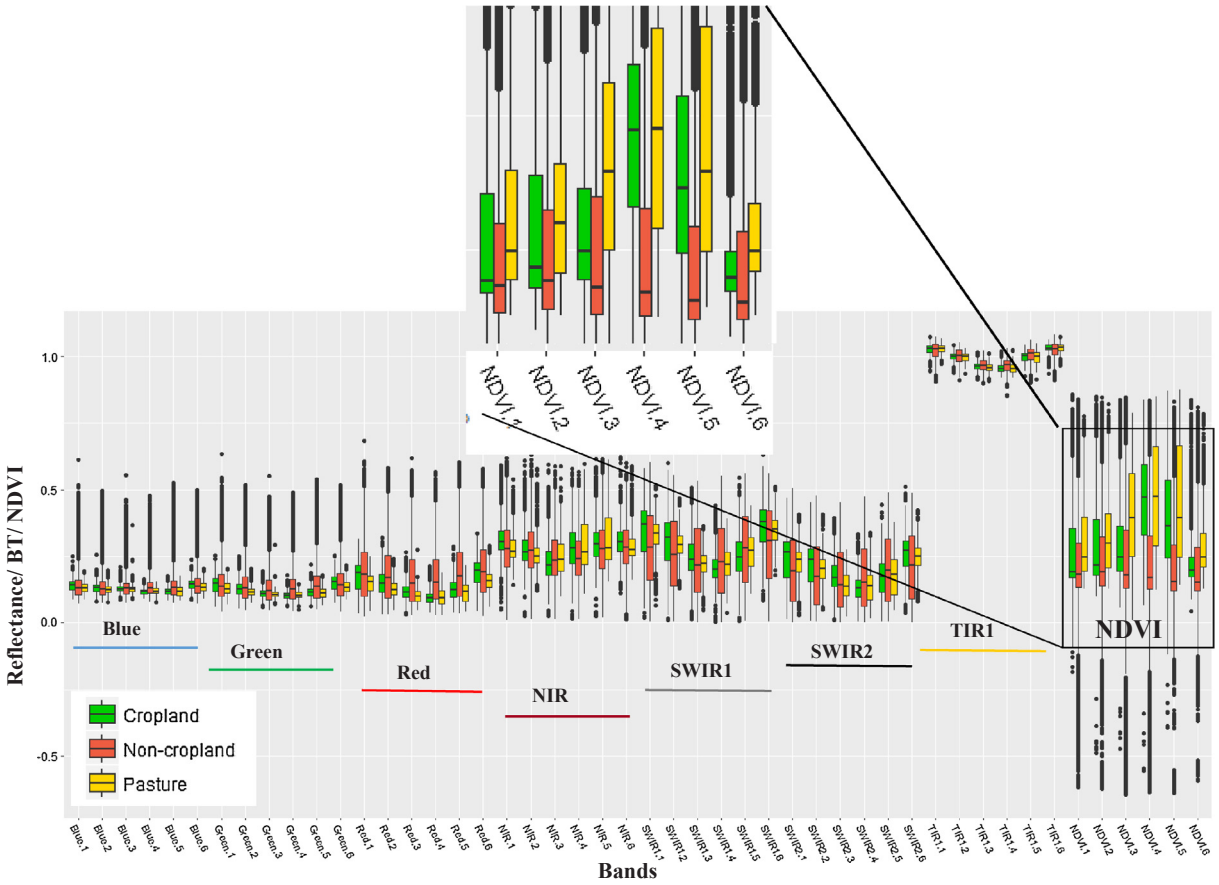

Fig. 6a. Croplands *versus* non-croplands knowledge capture to train random forest machine learning algorithm in Landsat bands for Australia of the study area (Band 1–6 (blue), band 7–12 (green), band 13–18 (NIR), band 19–24 (red), band 25–30 (SWIR1), 31–36 (SWIR2). 37–42 (TIR1), and 43–48 (NDVI). Note: TIR1 (thermal infrared band 1) values are scaled for convenience. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

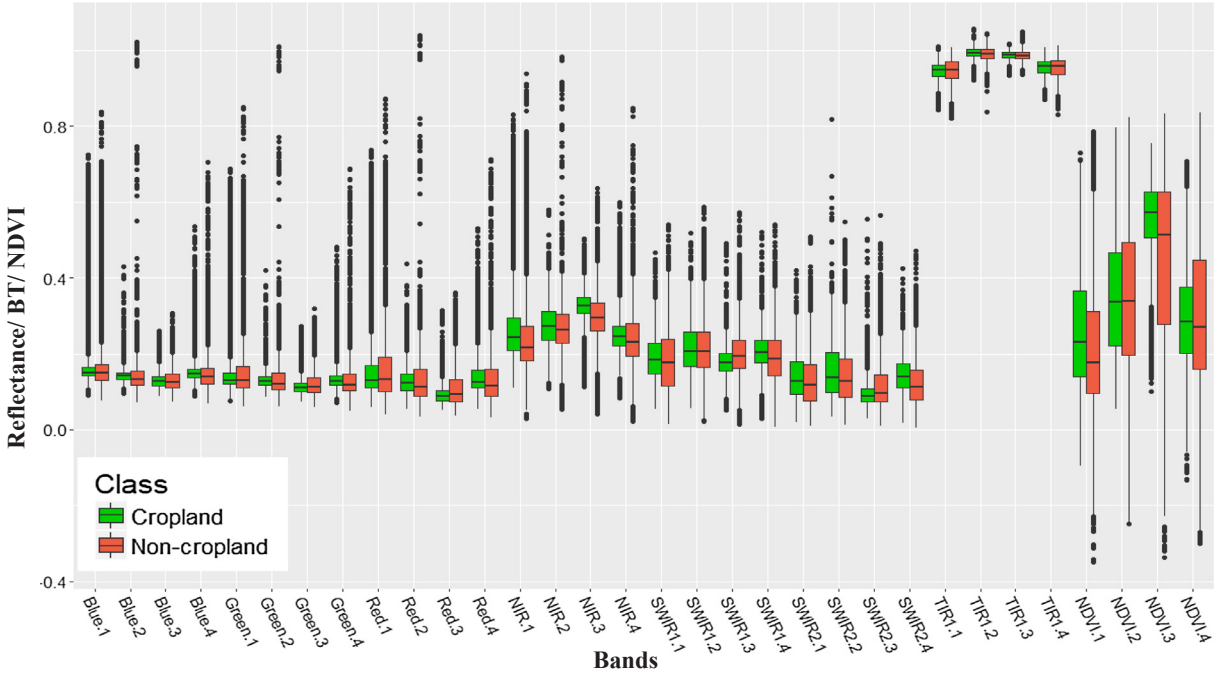

Fig. 6b. Croplands *versus* non-croplands knowledge capture to train random forest machine learning algorithm in Landsat bands for the Zone 2 (China) of the study area (Band 1–4 (blue), band 5–8 (green), band 9–12 (NIR), band 13–16 (red), band 17–20 (SWIR1), 21–24 (SWIR2). 25–28 (TIR1), and 29–32 (NDVI). Note: TIR1 (thermal infrared band 1) values are scaled for convenience. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with other cropland products generated with coarser spatial resolution data.

5.1. 30-m cropland extent product for Australia and China

The study produced a 30-m cropland extent map for the nominal year 2015 derived using Landsat 16-day time-series data for Australia (Fig. 7a) and China (Fig. 7b) by using the pixel-based supervised Random Forest (RF) machine learning algorithm for cropland classification. Initially, the RF machine learning algorithm, discussed in Section 4, were trained to separate croplands *versus* non-croplands for each of the zones (Fig. 1) based on knowledge generated using reference data (Section 3 and its sub-sections). The machine learning algorithm described in Section 4 and its sub-sections were then run on the GEE cloud-computing environment using Landsat data (Figs. 2a and 2b) for each of the zones to separate croplands *versus* non-croplands. The process was iterated and the training samples (e.g., Figs. 6a and 6b) in the algorithms were adjusted several times until we achieved optimal

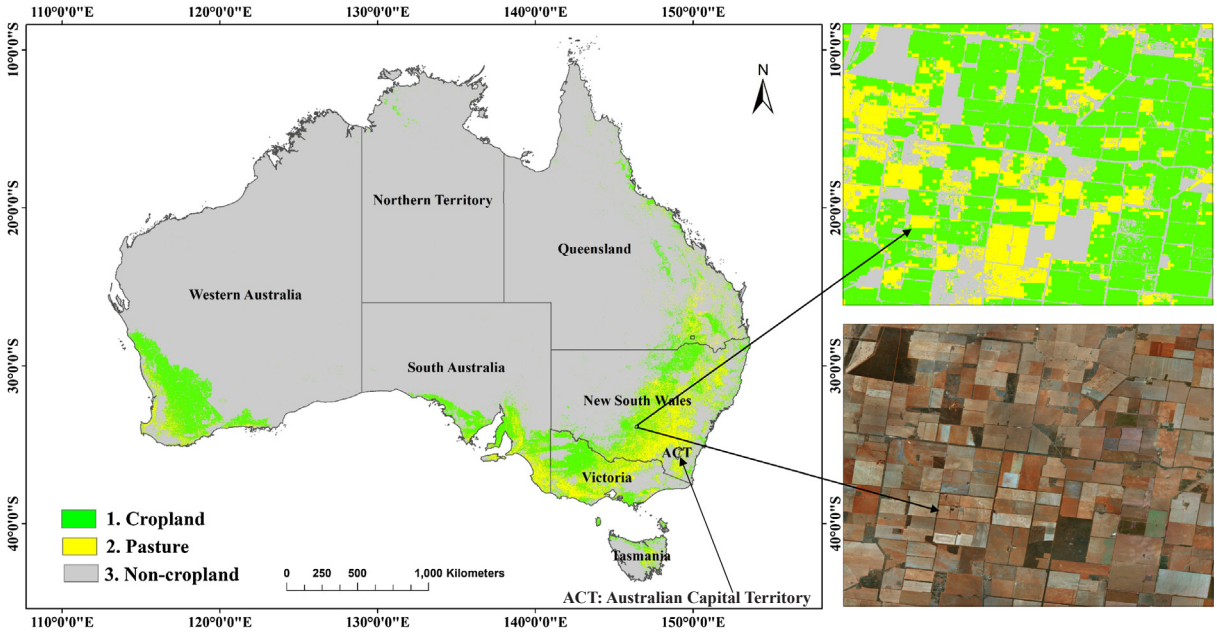

Fig. 7a. Cropland Extent Product at 30-m for Australia (left panel) with illustrative zoom in view for a location (right upper) and corresponding high-resolution imagery (right lower). This product is made available for visualization at www.croplands.org. The data are downloadable from LP DAAC: https://www.lpdaac.usgs.gov/node/1282.

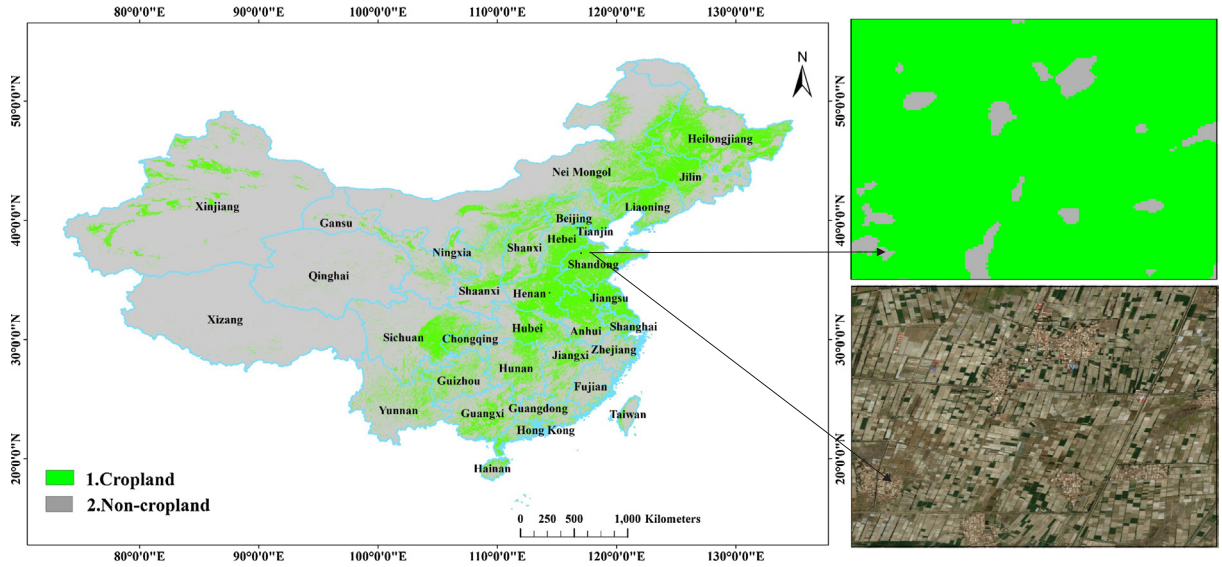

Fig. 7b. Cropland Extent Product at 30-m for China (left panel) with illustrative zoom in view for a location (right upper) and corresponding high-resolution imagery (right lower). This product is made available for visualization at www.croplands.org. The data are downloadable from LP DAAC: https://www.lpdaac.usgs.gov/node/1282.

results of croplands *versus* non-croplands (Figs. 7a and 7b below). The products are called the Global Food Security-support Analysis Data @ 30-m Cropland Extent for Australia (Fig. 7a), and China (Fig. 7b). This product is available through Land Processes Distributed Active Archive Center (LP DAAC https://www.lpdaac.usgs.gov/node/1282) for download (Teluguntla et al., 2017b) and the same dataset is also available for visualization at www.croplands.org.

5.2. Accuracy assessment

Accuracy assessment is a key component of map production using remote sensing data. The cropland extent maps for Australia, and China were assessed for their accuracies using error matrices with independent validation dataset (Congalton et al., 2017; Yadav and Congalton, 2017) that was not available to the producer of this dataset. Accuracy error matrices were produced separately for each of the zones as well as for the entire country.

Australia: A total of 900 stratified, randomly distributed validation samples were used to determine the accuracy of the final cropland extent map of Australia. An error matrix (Table 3) was generated for Australia providing producer's, user's, and overall accuracies (Congalton, 1991, 2009; Congalton and Green, 2008; Nhamo et al., 2018).

For the whole of Australia (zone# 1), the overall accuracy was 97.6% with producer's accuracy of 98.8% (errors of omissions = 1.2%) and user's accuracy of 79.0% (errors of commissions = 21%) (Table 3) for cropland class. These results clearly imply the high level of confidence in differentiating croplands from non-croplands for Australia. High producer's accuracy from the statistical assessment using stratified, randomly distributed samples across Australia (98.8%) suggest that almost all croplands of Australia were mapped accurately (with only 1.2% missing). However, as a trade-off, there is a commission error of 21% (meaning non-cropland mapped as croplands). When using any classification this trade-off is inevitable. If we reduce errors of commissions, errors of omissions will increase and vice versa. In this trade-off, it is better to capture almost all croplands even when we have some non-croplands also mapped as croplands. This provides a more complete cropland mask for higher level products.

Table 3
Accuracy error matrix. Accuracy assessment error matrix of the 30-m Cropland Extent Product for Australia.

Australia Zone# 1		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%**

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 4) were generated for each of the zones separately and for all of China providing producer's, user's and overall accuracies. For whole of China, the overall accuracy was 94% with producer's accuracy of 80% (errors of omissions = 20%) and user's accuracy of 84.2% (errors of commissions = 15.8%) (Table 4) for the cropland class. When considering all three zones (Table 4), 95.1% of total cropland area (TCA) of China is in zone #2 as shown in Table 4, 3.9% of total cropland area of China is in zone #3, and only 1% of total cropland area of China is in zone #4. Thereby, the accuracy in Zone #2 is the most important. This zone had very balanced and high producer's and user's accuracies with a producer's accuracy of 83.3% (errors of omission = 16.7%) and user's accuracy of 84.2% (errors of commission of 15.8%). The overall accuracies ranged between 91% and 98% for all zones. These results clearly imply the high level of confidence in differentiating croplands from non-croplands for China.

Within this project, the accuracy team conducted an assessment to determine the optimal number of samples required to perform accuracies. They concluded a total of 250 samples per zone as optimal, beyond which accuracies asymptote. Indeed, the total samples for training the classifiers and for accuracy assessments were significantly higher than the minimum of 250 samples. For example, accuracy assessments had: 900 samples for Australia, 1972 samples for China with the three zones in China having: 1184 for zone 2, 439 for zone 3, and 349 for zone 4. So, all zones have many more validation samples than the required optimal 250 samples.

It is possible to further increase accuracies of classification- especially the user's and producer's accuracies (Table 3 and 4) through number of measures. First, by breaking the zones (Fig. 1) to smaller ones; particularly zone 2 (eastern part of China) which contains approx.

Table 4
Accuracy error matrix. Independent accuracy assessment error matrix for the 30-m Cropland Extent Product for China by zone.

Zone 2

% of TCA*= 95.1%		Reference Data: Crop	Reference Data: 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 4

% of TCA*= 1.0%		Reference Data: Crop	Reference Data: 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 3

% of TCA*= 3.9%		Reference Data: Crop	Reference Data: 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 3 zones

		Reference Data: Crop	Reference Data: 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

95% of the total cropland area in China. This was not done in this study because of a requirement from the project independent accuracy team to minimize zones. Increasing the number of zones requires additional reference data to be acquired. This was not feasible within the scope of this project; however, is recommended in future work of this magnitude. Second, increased collection of reference training and validation data from agricultural fields of greater diversity (e.g., croplands along the mountains, flooded rice, fallows of various types, permanent crops of various types) is required to ensure that the classification algorithms are robust. Third, in addition to pixel-based RF MLA other pixel-based classifiers such as the support vector machines (SVMs) as well as the object-based algorithms such as the recursive hierarchical segmentation (RHSEG) will help resolve some of the uncertainties from classification results obtained from single algorithm (Xiong et al., 2017a). Fourth, the Google Earth Engine (GEE) is a powerful platform for image collection, organization, and classification over very large areas such as large countries, continents, and the entire planet through cloud computing platform. Whereas, GEE helps in powerful big-data management, and fast computing, there are several limitations on how the GEE processes the big-data using the MLAs. For example, for some classifiers, instead of classifying data of every pixel, it does "sampling" of the entire population of the big-data resulting in uncertainties in classification output and reduction in accuracies. There is also certain positional accuracy issue that sometimes can cause uncertainties on precise geographic location of a sample, especially when the pixel is 30-m resolution. To overcome this, we used a 90-m by 90-m sampling, yet, some uncertainties still need to be better understood. There are other ways of further improving accuracies and reducing uncertainties such as improved temporal resolution of imagery such as integration of global forest maps (Hansen et al., 2013) and global water mask (Carroll et al., 2009). Overall, Australia had nearly 2000 and China nearly 4000 training plus validation samples (Table 2). Given that the 30-m Landsat data over such large areas have billions of pixels, this is not a very large sample, but probably the best achievable given the complexity of the large areas and resources required. However, larger sample for both training and validation, especially from diverse croplands ranging from high mountains to lowland deltas that take into account various subclasses within standing crops, cropland fallows, and permanent crops will help produce products with even greater accuracies.

Similar methods to those used here were also applied to other continents of the world. For example, in a recent publication on Africa, Xiong et al. (2017a) reported a weighted overall accuracy of 94%, with a producer's accuracy of 85.9% (or omission error of 14.1%), and user's accuracy of 68.5% (commission error of 31.5%) for the cropland class across Africa. Xiong et al. (2017a) used two pixel-based machine learning algorithms (RF and SVM) along with an object-based classification (RHSEG). However, mapping croplands in Africa is much more complex than mapping croplands of Australia or China due to fragmented and mixed cropping in savannas and rainforests. As a result, the user's accuracies of Africa were relatively low compared to our study in Australia and China. Further, uncertainties in cropland training and validation samples is also somewhat higher in Africa than for Australia or China.

5.3. Cropland areas and comparsion with statistics from other sources

Apart from producing a map, calculation of cropland area statistics is an important component of the 30-m cropland extent product. Australia has two distinct cropland classes (croplands and managed pastures). Total cropland areas of the two classes (croplands + pasture) were estimated as 65.7 million hectares (Mha) of which croplands alone was 35.1 Mha for the year 2015. In contrast, cropland extent area of China was 165.2 Mha. Table 5 shows country-wise cropland area statistics generated from this study for Australia and China that were compared with other sources such as the national census data based MIRCA2005 (Stefan Siebert and Portmann, personal communication; Portmann et al., 2010) which was also updated in the year 2015 (Table 5) and national level official cropland statistics from Bureau of Statistics for China and Australia.

Table 5
Cropland area comparisons. Net cropland areas (NCAs) derived based on 30-m Global Food Security support Analysis Data (GFSAD30) cropland product and comparison with other cropland products.

Country	Land Area[a]	GFSAD30[b]	MIRCA 20,14[c]	National statistics[d,e]	% of Total Land Area
Name	Ha	Ha	Ha	Ha	%
Australia	768,851,504	35,105,792	30,615,114	32,078,000	4.6%
China	932,824,512	165,228,334	158,872,013	164,626,929	17.7%

[a] Total land area is land area excluding area under inland water bodies.
[b] GFSAD30 current study.
[c] Monthly irrigated and rainfed crop areas (MIRCA) around the year 2014 derived by Portmann et al. (2010).
[d] Australian Bureau of Statistics http://www.abs.gov.au/ausstats/abs@.nsf/mf/7121.0, total planted area for the year 2010–11, which is a good year in terms of climate.
[e] China statistical year book 2014 http://www.stats.gov.cn/tjsj/ndsj/2014/indexeh.htm.

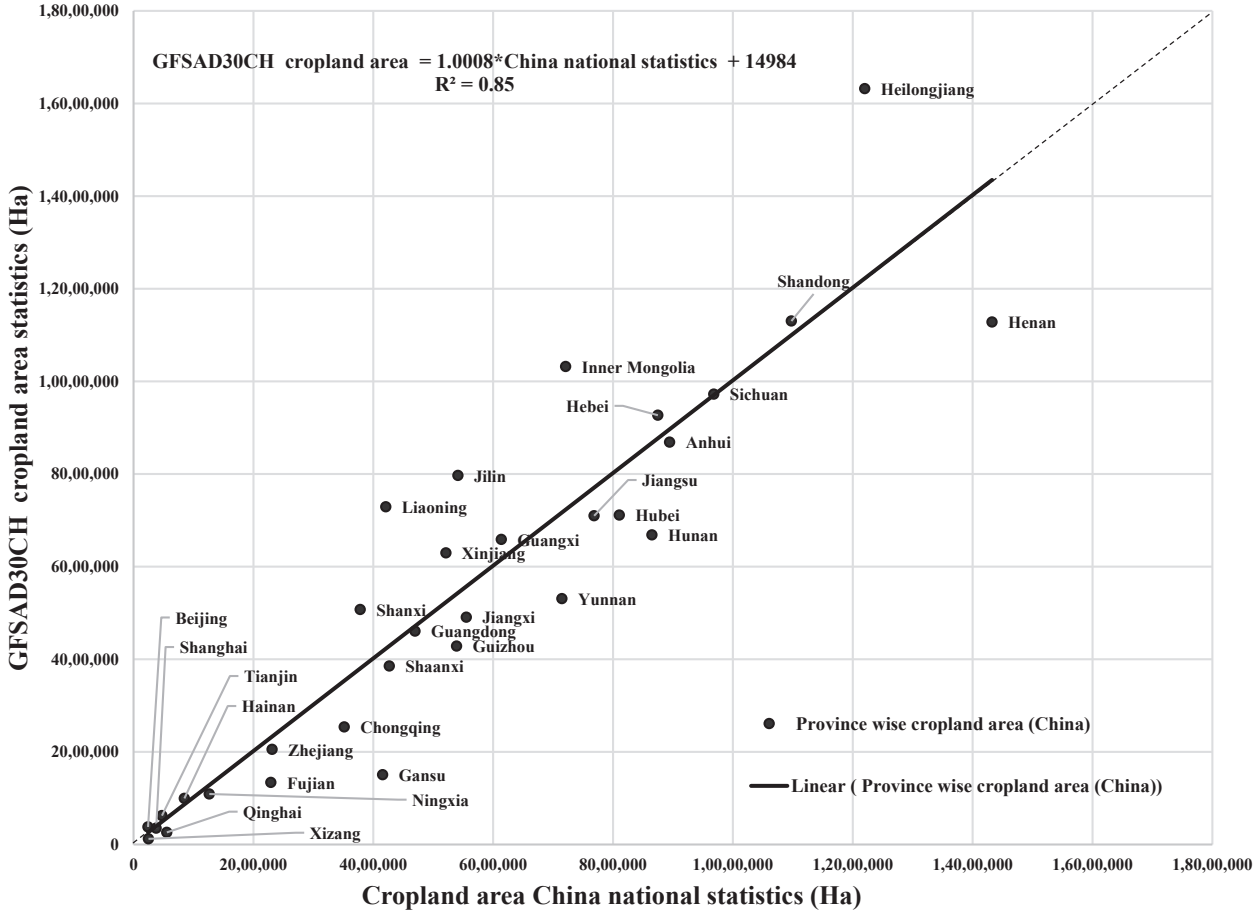

Fig. 8. Comparison with province-wise cropland areas. Comparison of province wise cropland areas derived from GFSAD30CE for China vs. cropland areas obtained from National statistics of China (http://www.stats.gov.cn).

However, 30-m cropland extent of Australia and China (www.croplands.org) produced by this study is at high resolution. Every 30-m pixel is 0.09 ha. Thereby, it is feasible to capture areas at the sub-national level, indeed even at individual farm level, small or big. This is a huge advantage over other existing cropland products. As a result, we calculated cropland areas of China at sub-national level for the 31 provinces of China and compared them with cropland areas from national statistics of China (http://www.stats.gov.cn/tjsj/ndsj/2014/indexeh.htm) as shown in Fig. 8. The relation between province wise areas from both sources show very high correlation with a slope value of 1.008 and R^2 value of 0.85. Among 31 provinces, two provinces Henan and Gansu shows lower-estimation relative to national statistics. This was due to some of the mountainous agriculture that was not completely captured by the RF algorithm in our study. Two provinces in the northeastern region of Heilongjiang, and Inner Mongolia shows higher-estimation in our study relative to national statistics (Fig. 8). This was due to some of the pasture which was classified as cropland in our cropland product.

5.4. Comparison of croplands mapped at different spatial scales/resolutions

Spatially, we have compared cropland products generated at three spatial resolutions with sub-meter to 5-m VHRI. The three-different spatial resolution cropland extent products are:

1. 30-m cropland extent product (this study),
2. 250-m cropland extent product (Teluguntla et al., 2017a), and
3. 1-km cropland extent product (GFSAD1-km, Teluguntla et al., 2015b)

Comparison of the above three products with VHRI (Fig. 9) showed the advantage of mapping croplands at 30-m finer spatial resolution (Fig. 9). We have presented 3 examples in Fig. 9. It is obvious that the field boundaries so well captured in 30-m cropland extent product (this study) are missing in 250-m and 1000-m. Further, omissions and commissions are high in coarser 250-m and 1-km products whereas the 30-m provides near precise estimates.

6. Conclusions

The study demonstrated a “paradigm-shift” in deriving high resolution 30-m cropland extent products over very large areas (e.g., Australia, China) using multi-year (2013–2015), 16-day Landsat data, random forest machine learning algorithms, large volume of reference training and validation datasets from multiple sources by utilizing the big-data management and processing power of Google Earth Engine cloud-computing platform. These methods and approaches demonstrated the ability to map croplands rapidly and accurately at continental\big-country scale involving petabyte volume big-data. The resulting 30-m Landsat-derived cropland extent products of Australia and China for the nominal year 2015 had overall accuracies of over 94%. The cropland class was mapped with producer’s accuracy of 98.8% (errors of omissions = 1.2%) for Australia and 80% (errors of omissions = 20%) for China. The user’s accuracies were 79% (errors of commissions = 21%) for Australia, and 84.2% (errors of commissions = 15.8%) for China. The study also demonstrated the ability to compute sub-national province level cropland area statistics that explained 85% variability when compared with national statistics of China. The 30-m cropland extent product mapped complex series of

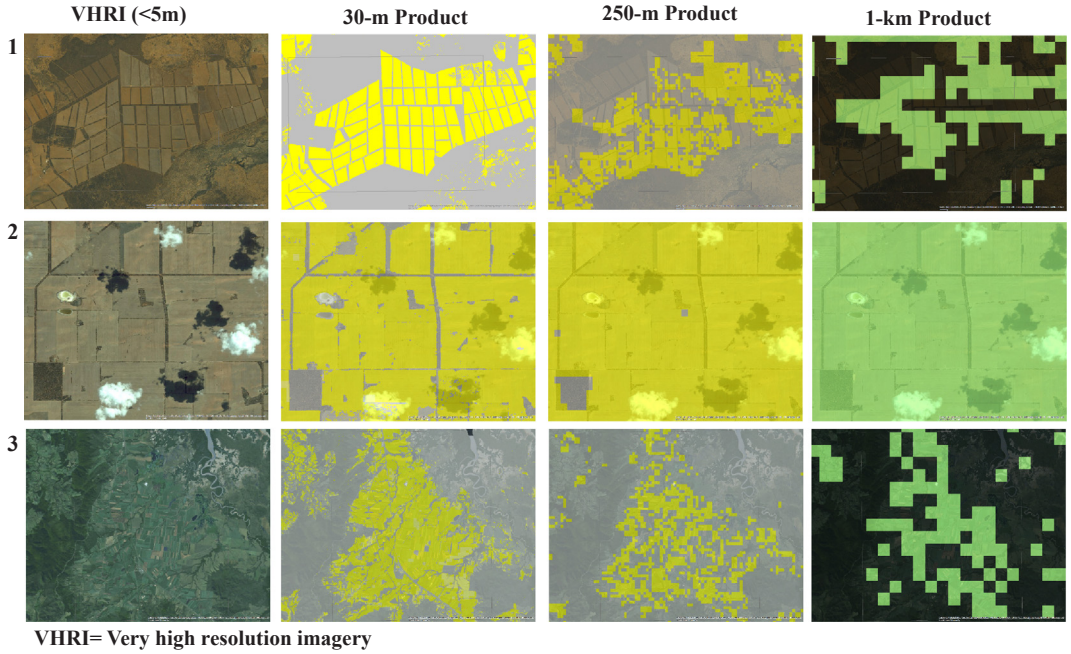

Fig. 9. Spatial comparison of cropland products for precision. Comparison of cropland extent products across three different spatial resolutions (30-m, 250-m, and 1000-m) with < 5-m very high-resolution imagery.

small and large farms of entire Australia and China accurately and precisely and can be viewed at full resolution by browsing in www.croplands.org. The products are made available for download at NASA's Land Processes Distributed Active Archive Center (LPDAAC): https://www.lpdaac.usgs.gov/node/1282.

Acknowledgements

Authors are grateful for the funding received through NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs), through NASA ROSES solicitation (June 1, 2013–May 31, 2018). This funding was received through NASA MEaSUREs project grant number: NNH13AV82I and the USGS Sales Order number 29039. The United States Geological Survey (USGS) provided significant direct and indirect supplemental funding through its Land Resources Mission Area (LRMA), National Land Imaging Program (NLIP), and Land Change Science (LCS) program. We gratefully acknowledge this. This research is a part of the Global Food Security -support Analysis Data Project at 30-m (GFSAD30). Authors would like to thank the Google Earth Engine program for providing computing resources and coding support. The authors would like to thank Dr. Felix T. Portman and Dr. Stefan Siebert for providing statistics of MIRCA2000 (Portmann et al., 2010; also, latest statistics through personal communication between. Dr. Stefan Siebert and Prasad S. Thenkabail) for inter-comparison. The authors would like to thank Dr. Rakhesh Devadas, and two student volunteers, University of Technology Sydney, Sydney for the comprehensive fieldwork support and coordination in Australia. Special thanks to Dr. Itiya Aneece and Dr. Naga Manohar Velpuri for their suggestions to the manuscript. We would like to thank the anonymous reviewers that helped improve this paper with their constructive comments and suggestions. Finally, special thanks to Ms. Susan Benjamin, Director of Western Geographic Science Center (WGSC) of USGS, Mr. Larry Gaffney, Administrative officer of WGSC of USGS, and Mr. Mark Sittloh, CEO of the Bay Area Environmental Research Institute (BAERI) for their support and encouragement throughout this research.

References

ABARES, 2016. Australian Crop Report Series, Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra, February. <http://www.agriculture.gov.au/abares> (accessed March 2016).
Arsanjani, J.J., Tayyebi, A., Vaz, E., 2016. GlobeLand30 as an alternative fine-scale global land cover map: challenges, possibilities, and implications for developing countries. Habitat Int. 55, 25–31.
Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.
Biggs, T., Thenkabail, P., Gumma, M., Scott, C., Parthasaradhi, G., Turral, H., 2006. Irrigated area mapping in heterogeneous landscapes with MODIS time series, ground truth and census data, Krishna Basin, India. Int. J. Remote Sens. 27, 4245–4426.
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. Int. J. Appl. Earth Observ. Geoinformat. 11, 114–129.
Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. Geocarto Int. 26 (5), 341–358.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Carroll, M.L., Townshend, J.R., DiMiceli, C.M., Noojipady, P., Sohlberg, R.A., 2009. A new global raster water mask at 250 m resolution. Int. J. Digital Earth 2 (4), 291–308.
Chan, J.C.W., Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sens. Environ. 112 (6), 2999–3011.
Chander, G., Markham, B.L., Helder, D.L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sens. Environ. 113, 893–903.
Chen, C., Liaw, A., Breiman, L., 2004. Using Random Forest to Learn Imbalanced Data, Technical Report 666. . University of California, Berkeley.
Chen, G., Thill, J.C., Anantsuksomsri, S., Tontisirin, N., Tao, R., 2018. Stand age estimation of rubber (*Hevea brasiliensis*) plantations using an integrated pixel-and object-based tree growth model and annual Landsat time series. ISPRS J. Photogramm. Remote Sens. 144, 94–104.
Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., 2015. Global land cover mapping at 30m resolution: a POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7–27.
Cohen, W.B., Goward, S.N., 2004. Landsat's role in ecological applications of remote sensing. Bioscience 54, 535–545.
Congalton, R., 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 583–602.
Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46.
Congalton, R.G., 2009. 19 accuracy and error analysis of global and local maps: lessons learned and future considerations. Remote Sens. Glob. Croplands Food Secur. 441.
Congalton, R.G., Green, K., 2008. Assessing the Accuracy of Remotely Sensed Data:

Principles and Practices. CRC Press.
Congalton, G., R., Yadav, K., McDonnell, K., Poehnelt, J., Stevens, B., Gumma, K.M., Teluguntla, P., Thenkabail, P.S., 2017. NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) Cropland Extent 2015 Validation Global 30 m V001 [Data set]. NASA EOSDIS Land Processes DAAC. http://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30VAL.001.
Defourny P., Schouten, L., Bartalev, S., Bontemps, S., Caccetta, P., de Witt, A., di Bella, C., Gerard, B., Giri, C., Gond, V., Hazeu, G., Heinimann, A., Herold, M., Jaffrain, G., Latifovic, R., Ling, H., Mayaux, P., Muncher, S., Nonguierma, A., Stibig, H.-J., Van Bogaert, E., Vancutsem, C., Bicheron, P., Leroy, M., Arino, O., 2009. Accuracy assessment of a 300-m global land cover map: the GlobCover experience. In: 33rd International Symposium on Remote Sensing of Environment (ISRSE), Stresa, Italy, May 2009.
DeFries, R., Chan, J.C.-W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sens. Environ. 74, 503–515.
Dheeravath, V., Thenkabail, P.S., Chandrakantha, G., Noojipady, P., Reddy, G.P.O., Biradar, C.M., Velpuri, M., et al., 2010. Irrigated areas of India derived using MODIS 500 m time series for the years 2001–2003. ISPRS J. Photogramm. Remote Sens. 65 (1), 42–59.
Duro, D.C., Franklin, S.E., Dubé, M.G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens. Environ. 118, 259–272.
Erickson, T., 2014. Multi-source Geospatial Data Analysis with Google Earth Engine, Fall Meeting 2014, abstract #IN53E-05. American Geophysical Union.
Fischer, G., van Velthuizen, H., Nachtergaele, F.O., 2000. Global Agroecological Zones Assessment: Methodology and Results (IIASA Interim Report IR-00-064).
Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., 2011. Solutions for a cultivated planet. Nature 478, 337–342.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.
Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Havlik, P., et al., 2015. Mapping global cropland and field size. Global change biology 21 (5), 1980–1992.
Funk, C.C., Brown, M.E., 2009. Declining global per capita agricultural production and warming oceans threaten food security. Food Secur. 1, 271–289.
Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27 (4), 294–300.
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27.
Gumma, M.K., Nelson, A., Thenkabail, P.S., Singh, A.N., 2011. Mapping rice areas of South Asia using MODIS multitemporal data. J. Appl. Remote Sens. 5 053547-053547-053526.
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. Int. J. Digital Earth 9, 981–1003.
Gumma, M.K., Thenkabail, P.S., Maunahan, A., Islam, S., Nelson, A., 2014. Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500 m data for the year 2010. ISPRS J. Photogramm. Remote Sens. 91, 98–113.
Gumma, M.K., Thenkabail, P.S., Deevi, K.C., Mohammed, I.A., Teluguntla, P., Oliphant, A., Whitbread, A.M., et al., 2018. Mapping cropland fallow areas in myanmar to scale up sustainable intensification of pulse crops in the farming system. GISci. Remote Sens. 1–24.
Gutman, G., Byrnes, R., Masek, I., Covington, S., Justice, C., Franks, S., Headley, R., 2008. Towards monitoring changes at a Globa the Global Land S. Photogramm. Eng. Remote Sens. 74, 6–10.
Han, W., Yang, Z., Di, L., Mueller, R., 2012. CropScape: a Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. Comput. Electron. Agric. 84, 111–123.
Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Kommareddy, A., et al., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850–853.
Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: the Landsat data continuity mission. Remote Sens. Environ. 122, 11–21.
Johnson, D.M., Mueller, R., 2010. The 2009 cropland data layer. PE&RS Photogramm. Eng. Remote Sens. 76 (11), 1201–1205.
Kumar, A., Hubert-Moy, L., Buvaneshwari, S., Sekhar, M., Ruiz, L., Bandyopadhyay, S., Corgne, S., 2018. Irrigation history estimation using multitemporal Landsat satellite Images: application to an intensive groundwater irrigated agricultural watershed in India. Remote Sens. 10 (6), 893.
Landsat 8 OLI and TIRS. Available online: https://www.lta.cr.usgs.gov/L8 (accessed on 4 December 2017).
Lawrence, R.L., Wood, S.D., Sheley, R.L., 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). Remote Sens. Environ. 100, 356–362.
Liang, D., Zuo, Y., Huang, L., Zhao, J., Teng, L., Yang, F., 2015. Evaluation of the consistency of MODIS land cover product (MCD12Q1) based on Chinese 30 m globeland30 datasets: a case study in Anhui Province, China. ISPRS Int. J. Geo-Inf. 4, 2519–2541.
Loveland, T., Reed, B., Brown, J., Ohlen, D., Zhu, Z., Yang, L., Merchant, J., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sens. 21, 1303–1330.
Lymburner, L., Tan, P., Mueller, N., Thackway, R., Lewis, A., Thankappan, M., Randall, L., Islam, A., Senarath, U., 2011. The national dynamic land cover dataset (DLCD), record 2011/31. In.
Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.-K., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. Geosci. Remote Sens. Lett. IEEE 3, 68–72.
Marshall, M., Thenkabail, P., 2015. Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and MODIS vegetation indices in crop biomass estimation. ISPRS J. Photogramm. Remote Sens. 108, 205–218.
Matejicek, L., Kopackova, V., 2010. Changes in croplands as a result of large scale mining and the associated impact on food security studied using time-series Landsat images. Remote Sens. 2 (6), 1463–1480.
Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 66, 247–259.
Na, X., Zhang, S., Li, X., Yu, H., Liu, C., 2010. Improved land cover mapping using random forests combined with landsat thematic mapper imagery and ancillary geographic data. Photogramm. Eng. Remote Sens. 76, 833–840.
Nhamo, L., van Dijk, R., Magidi, J., Wiberg, D., Tshikolomo, K., 2018. Improving the accuracy of remotely sensed irrigated areas using post-classification enhancement through UAV capability. Remote Sens. 10 (5), 712.
Ozdogan, M., Gutman, G., 2008. A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US. Remote Sens. Environ. 112 (9), 3520–3537.
Ozdogan, M., Woodcock, C.E., 2006. Resolution dependent errors in remote sensing of cultivated areas. Remote Sens. Environ. 103, 203–217.
Pan, Z., Huang, J., Zhou, Q., Wang, L., Cheng, Y., Zhang, H., Blackburn, G.A., Yan, J., Liu, J., 2015. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. Int. J. Appl. Earth Observ. Geoinformat. 34, 188–197.
Pantazi, X., Moshou, D., Alexandridis, T., Whetton, R., Mouazen, A., 2016. Wheat yield prediction using machine learning and advanced sensing techniques. Comput. Electron. Agric. 121, 57–65.
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 Sens. Environ. 187, 156–168.
Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115 (6), 1301–1316.
Peña, J.M., Gutiérrez, P.A., Hervás-Martínez, C., Six, J., Plant, R.E., López-Granados, F., 2014. Object-based image classification of summer crops with machine learning methods. Remote Sens. 6 (6), 5019–5041.
Pervez, M.S., Budde, M., Rowland, J., 2014. Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI. Remote Sens. Environ. 149, 155–165.
Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., Justice, C.O., 2010. Estimating global cropland extent with multi-year MODIS data. Remote Sensing 2, 1844–1863.
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. Glob. Biogeochem. Cycles 24.
Roy, D.P., Wulder, M., Loveland, T.R., Woodcock, C., Allen, R., Anderson, M., Helder, D., Irons, J., Johnson, D., Kennedy, R., 2014. Landsat-8: science and product vision for terrestrial global change research. Remote Sens. Environ. 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. Int. J. Appl. Earth Observ. Geoinformat. 38, 321–334.
Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M., West, P.C., 2016. Subnational distribution of average farm size and smallholder contributions to global food production. Environ. Res. Lett. 11, 124010.
Shao, Y., Lunetta, R.S., 2012. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS J. Photogramm. Remote Sens. 70, 78–87.
Siebert, S., Döll, P., 2010. Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. J. Hydrol. 384, 198–217.
Sun, Y., Kamel, M.S., Wong, A.K., Wang, Y., 2007. Cost-sensitive boosting for classification of imbalanced data. Pattern Recogn. 40, 3358–3378.
Tatsumi, K., Yamashiki, Y., Torres, M.A.C., Taipe, C.L.R., 2015. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. Comput. Electron. Agric. 115, 171–179.
Teluguntla, P., Ryu, D., George, B., Walker, J.P., Malano, H.M., 2015a. Mapping flooded rice paddies using time series of MODIS imagery in the Krishna River Basin, India. Remote Sens. 7, 8858–8888.
Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Giri, C., Milesi, C., Ozdogan, M., Congalton, R.G., Tilton, J., Sankey, T.T., 2015b. 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. Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. CRC Press, pp. 131–159.
Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Congalton, R.G., Oliphant, A., Poehnelt, J., Yadav, K., Rao, M., Massey, R., 2017a. Spectral matching techniques

(SMTs) and automated cropland classification algorithms (ACCAs) for mapping croplands of Australia using MODIS 250-m time-series (2000–2015) data. Int. J. Digital Earth 1–34.

Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., G., Congalton, R., Oliphant, J., A., Sankey, T., Poehnelt, J., Yadav, K., Massey, R., Phalke, A., Smith, C., 2017b. 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. http://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD30AUNZCNMOCE.001.

Thenkabail, P., GangadharaRao, P., Biggs, T., Krishna, M., Turral, H., 2007. Spectral matching techniques to determine historical land-use/land-cover (LULC) and irrigated areas using time-series 0.1-degree AVHRR Pathfinder datasets. Photogramm. Eng. Remote Sens. 73, 1029–1040.

Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., Cai, X., 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. Int. J. Remote Sens. 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 Sens. 2 (1), 211–261.

Thenkabail, P.S., Knox, J.W., Ozdogan, M., Gumma, M.K., Congalton, R.G., Wu, Z., Milesi, C., Finkral, A., Marshall, M., Mariotto, I., 2012. Assessing future risks to agricultural productivity, water resources and food security: how can remote sensing help? Photogramm. Eng. Remote Sens. 78, 773–782.

Thenkabail, P.S., Wu, Z., 2012. An automated cropland classification algorithm (ACCA) for Tajikistan by combining Landsat, MODIS, and secondary data. Remote Sens. 4, 2890–2918.

Tilton, J.C., Tarabalka, Y., Montesano, P.M., Gofman, E., 2012. Best merge region-growing segmentation with integrated nonadjacent region object aggregation. IEEE Trans. Geosc. Remote Sens. 50, 4454–4467.

Velpuri, N.M., Thenkabail, P.S., Gumma, M.K., Biradar, C., Dheeravath, V., Noojipady, P., Yuanjie, L., 2009. Influence of resolution in irrigated area mapping and area estimation. Photogramm. Eng. Remote Sens. 75 (12), 1383–1395.

Verburg, P.H., Mertz, O., Erb, K.-H., Haberl, H., Wu, W., 2013. Land system change and food security: towards multi-scale land system solutions. Curr. Opin. Environ. Sustain. 5, 494–502.

Waldner, F., Canto, G.S., Defourny, P., 2015. Automated annual cropland mapping using knowledge-based temporal features. ISPRS J. Photogramm. Remote Sens. 110, 1–13.

Waldner, F., Fritz, S., Di Gregorio, A., Plotnikov, D., Bartalev, S., Kussul, N., Gong, P., Thenkabail, P., Hazeu, G., Klein, I., 2016. A unified cropland layer at 250 m for global agriculture monitoring. Data 1, 3.

Wang, J., Zhao, Y., Li, C., Yu, L., Liu, D., Gong, P., 2015. Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. ISPRS J. Photogramm. Remote Sens. 103, 38–47.

Xiao, X., Boles, S., Frolking, S., Li, C., Babu, J.Y., Salas, W., Moore, B., 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. Remote Sens. Environ. 100, 95–113.

Xiong, J., Thenkabail, P., Tilton, J., Gumma, M., Teluguntla, P., Oliphant, A., Congalton, R., Yadav, K., Gorelick, N., 2017a. Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on Google Earth Engine. Remote Sens. 9, 1065.

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., Thau, D., 2017b. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS J. Photogramm. Remote Sens. 126, 225–244.

Yadav, K., Congalton, R.G., 2017. Issues with large area thematic accuracy assessment for mapping cropland extent: a tale of three continents. Remote Sens. 10 (1), 53.

Yang, Y., Xiao, P., Feng, X., Li, H., 2017. Accuracy assessment of seven global land cover datasets over China. ISPRS J. Photogramm. Remote Sens. 125, 156–173.

Yu, L., Wang, J., Clinton, N., Xin, Q., Zhong, L., Chen, Y., Gong, P., 2013. FROM-GC: 30 m global cropland extent derived through multisource data integration. Int. J. Digital Earth 6, 521–533.

Zhong, L., Hu, L., Yu, L., Gong, P., Biging, G.S., 2016. Automated mapping of soybean and corn using phenology. ISPRS J. Photogramm. Remote Sens. 119, 151–164.

Zhou, Y., Xiao, X., Qin, Y., Dong, J., Zhang, G., Kou, W., Jin, C., Wang, J., Li, X., 2016. Mapping paddy rice planting area in rice-wetland coexistent areas through analysis of Landsat 8 OLI and MODIS images. Int. J. Appl. Earth Observ. Geoinformat. 46, 1–12.