Remote Sensing of Environment 190 (2017) 383–395

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey

Xiao-Peng Song [a,*], Peter V. Potapov [a], Alexander Krylov [a], LeeAnn King [a], Carlos M. Di Bella [b,c,d], Amy Hudson [a], Ahmad Khan [a], Bernard Adusei [a], Stephen V. Stehman [e], Matthew C. Hansen [a]

[a] Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA
[b] INTA-Castelar, CNIA, Instituto de Clima y Agua, Las Cabañas y Los Reseros S/N, Hurlingham 1686, Buenos Aires, Argentina
[c] CONICET - Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina
[d] Departamento de Métodos Cuantitativos, Facultad de Agronomía, Universidad de Buenos Aires, Argentina
[e] College of Environmental Science and Forestry, State University of New York, Syracuse, NY 13210, USA

ARTICLE INFO

Article history:
Received 24 May 2016
Received in revised form 9 October 2016
Accepted 8 January 2017
Available online 16 January 2017

Keywords:
Agriculture
Cropland
Sample
Remote sensing
Landsat
Image time-series
Classification
Decision tree

ABSTRACT

Reliable and timely information on agricultural production is essential for ensuring world food security. Freely available medium-resolution satellite data (e.g. Landsat, Sentinel) offer the possibility of improved global agriculture monitoring. Here we develop and test a method for estimating in-season crop acreage using a probability sample of field visits and producing wall-to-wall crop type maps at national scales. The method is illustrated for soybean cultivated area in the US for 2015. A stratified, two-stage cluster sampling design was used to collect field data to estimate national soybean area. The field-based estimate employed historical soybean extent maps from the U.S. Department of Agriculture (USDA) Cropland Data Layer to delineate and stratify U.S. soybean growing regions. The estimated 2015 U.S. soybean cultivated area based on the field sample was 341,000 km^2 with a standard error of 23,000 km^2. This result is 1.0% lower than USDA's 2015 June survey estimate and 1.9% higher than USDA's 2016 January estimate. Our area estimate was derived in early September, about 2 months ahead of harvest. To map soybean cover, the Landsat image archive for the year 2015 growing season was processed using an active learning approach. Overall accuracy of the soybean map was 84%. The field-based sample estimated area was then used to calibrate the map such that the soybean acreage of the map derived through pixel counting matched the sample-based area estimate. The strength of the sample-based area estimation lies in the stratified design that takes advantage of the spatially explicit cropland layers to construct the strata. The success of the mapping was built upon an automated system which transforms Landsat images into standardized time-series metrics. The developed method produces reliable and timely information on soybean area in a cost-effective way and could be applied to other regions and potentially other crops in an operational mode.

1. Introduction

Reliable and timely information on agricultural production is essential for ensuring world food security. Traditionally, agricultural data are acquired through census and ground survey. While ground-based data collection has the advantage of obtaining a wide range of variables related to the organizational structure of agriculture, such as land tenure, farm size, labor, crop area, irrigation, and fertilizer use, agricultural censuses are usually undertaken at a decadal frequency and thus they are most suitable to represent those aspects of agriculture that change slowly over time (FAO, 2015). Using census data across the globe would also encounter the data inconsistency problem, including inconsistent definitions of census variables, changing political or sampling units and various reporting protocols among different countries and census intervals (Portmann et al., 2010; Ramankutty et al., 2008).

Satellite observations, owing to their synoptic and repetitive nature, have the unique advantage of providing timely and spatially contiguous information on crop growth at regional to global scales. However, identification of crop type using satellite data remains a technical challenge due to the diversity of cropping systems, including crop types, crop varieties, management practices and field sizes. Thus, global cropland monitoring requires data of high spatial and temporal resolutions (Cihlar, 2000; Fritz et al., 2015; Thenkabail et al., 2010; Waldner et al., 2016). To date, satellite data of fine temporal resolution and coarse spatial resolution are predominantly used in agricultural research, especially over large areas. For example, data from the Moderate Resolution Imaging Spectroradiometer (MODIS) have been extensively used in cropland mapping (e.g. Chang et al., 2007; Lobell and Asner, 2004; Ozdogan, 2010; Wardlow and Egbert, 2008), as well as yield estimation

* Corresponding author.
E-mail address: xpsong@umd.edu (X.-P. Song).

http://dx.doi.org/10.1016/j.rse.2017.01.008
0034-4257/

(e.g. Anderson et al., 2016; Becker-Reshef et al., 2010; Bolton and Friedl, 2013; Doraiswamy et al., 2005; Johnson, 2014; Lopresti et al., 2015). At the global scale, cropland is often characterized as one or a few aggregated land cover classes at moderate to coarse resolutions (30 m–1 km) (e.g. Friedl et al., 2002; Gong et al., 2013; Hansen et al., 2000). Crop-specific masks are available at even coarser resolutions (~10 km) (Portmann et al., 2010; Ramankutty et al., 2008; Thenkabail et al., 2009; You et al., 2014). Since the opening of the Landsat archive in year 2008, the recent launch of Landsat 8 and Sentinel-2, medium-resolution data have shown great potential of producing rich temporal information that is previously available only with coarse-resolution data (Hansen et al., 2014; Roy et al., 2010). Studies have begun to take this opportunity in research on crop classification (e.g. Zhong et al., 2014) and yield estimation (e.g. Lobell et al., 2015) albeit over small areas.

The idea of using satellite data in operational agricultural surveys was formulated in the 1970s and exploratory programs were soon initiated, such as the Large Area Crop Inventory Experiment (LACIE) (Macdonald and Hall, 1980), Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing (AgRISTARS) and Monitoring Agriculture with Remote Sensing (MARS) (https://ec.europa.eu/jrc/en/mars). Today, two prominent examples of operational use of medium-resolution satellite data for national-scale crop type mapping are the Cropland Data Layer (CDL) generated by the United States Department of Agriculture (USDA) (Johnson and Mueller, 2010) and the Crop Inventory (CI) dataset generated by the Agriculture and Agri-Food Canada (AAFC). CDL has been produced annually since 2008 and CI has been produced annually since 2011, both with national coverage and at 30–56 m of spatial resolutions. Both CDL and CI are generated using supervised classification approaches and rely on comprehensive survey-based geospatial data for training. The sheer volume of the training data contributes to producing highly accurate maps, both over 85% for major crops in major agricultural states (AAFC, 2015; Boryan et al., 2011). However, not only are these training datasets not publically available, but they are also financially expensive and time-consuming to generate and update every year. Such datasets or the capacity to generate them may not be readily available in other countries, especially developing countries. Therefore, the methodology for producing CDL or CI is difficult to implement elsewhere.

Crop classification maps of high overall accuracy such as CDL and CI cannot be directly used through "pixel counting" for acreage estimation, because map products are usually biased due to misclassification and the existence of mixed pixels (Gallego, 2004). One way of deriving crop acreage and uncertainty estimates is to use a probability sample in an area sampling frame (AFS) (Carfagna and Gallego, 2005; Cotter and Tomczack, 1994; Pradhan, 2001). The classification map can be used as an ancillary variable and survey estimates as the dependent variable to perform regression analysis (Boryan et al., 2011; Gallego et al., 2014; Gonzáles-Alonso and Cuevas, 1993; Gonzáles-Alonso et al., 1991; Hill and Megier, 1988). Map-making and sample-based area estimation are often carried out as independent processes. But opportunities exist to reconcile the discrepancy of map-based and sample-based area estimates by closely integrating the two processes. Such a general approach has been successfully applied for mapping and estimating areas of forest cover change at continental to global scales (Broich et al., 2011; Hansen et al., 2010; Hansen et al., 2008b; Tyukavina et al., 2015).

The objective of this study is to develop a method applicable at national scales for estimating in-season cultivated area for a specific crop as well as to produce a spatially explicit crop cover map. Specifically, we estimate soybean cultivation area in the United States in year 2015 using a probability sample of field visits and we map soybean cover using all available Landsat data in the growing season of the year 2015. Soybean in the U.S. is chosen not only because the U.S. is the world's leading producer of this major commodity crop but also because independent data exist to provide a comparison with our results. The developed procedure is expected to be applicable to other crops and to other regions such as Brazil and Argentina where industrial monoculture dominates agricultural production.

2. Data and methods

2.1. Sample-based soybean area estimation

2.1.1. Study area and sampling design

We implemented a two-stage cluster sampling design for estimating national soybean area within the U.S. In the first stage, 20 km × 20 km blocks (clusters) were selected using a stratified random sampling technique. In the second stage, 30 m × 30 m Landsat pixels were selected using simple random sampling. The large agricultural fields in the U.S.—mean size 0.193 km^2 and median size 0.278 km^2 (Yan and Roy, 2016), ensured that our 20 km × 20 km blocks contained a large number of fields whereas the majority of our 30 m × 30 m pixels were pure pixels of a single crop. The cluster design was chosen to reduce the cost of field visits by spatially constraining the sample pixels to the selected clusters. The specific size of the cluster (20 km × 20 km) was chosen because it allowed the second-stage sample for each cluster to be completed in a single day.

The entire conterminous U.S. land area was divided into a regular grid of 20 km × 20 km blocks consisting of 20,371 blocks. We acquired the 30 m spatial resolution CDL for years between 2010 and 2014 and calculated the 5-year average soybean percentage for every block. For each block in each year, we counted the number of pixels labeled as soybean and divided the number by the total number of pixels within the block to compute the soybean percentage at the 20 km resolution. We then computed the arithmetic mean of soybean percentage over 2010–2014 for every block. After sorting the blocks based on the 2010–2014 mean soybean percentage from the largest to the smallest, the blocks that cumulatively accounted for 99.9% national soybean area were selected to create the fixed population ($N = 7028$ blocks).

Two-stage stratified random sampling was implemented, where the population was divided into four strata according to soybean intensity and a total of 70 sample blocks were randomly selected (Fig. 1, Table 1). The strata definitions and sample sizes were guided by evaluating the precision for estimating soybean area assuming the CDL soybean area was the population of interest. That is, we used the CDL soybean to obtain per-stratum means and variances and this allowed us to compute the standard error of estimated soybean area for various choices of strata and sample size allocation to strata (see Section 2.1.3 for estimation formulas). Based on these analyses, we determined that a sample size of 70 blocks was affordable while still likely to yield estimates that would have adequate precision to demonstrate the utility of the approach.

The second-stage sampling was constrained to the cropland region within each block to avoid visiting remote non-cropland sites, assuming no significant land use conversion between cropland and non-cropland within a year. We created a 2010–2014 maximum cropland mask for each block. A 30 m × 30 m pixel was included in this cropland mask if it was classified by CDL as cropland in any year between 2010 and 2014. Using simple random sampling, we selected 10 pixels within the cropland mask of each first-stage sample block. A total of 700 pixels were selected for field visit, 200 from each of the high, medium, and low strata, and 100 from the very low stratum.

2.1.2. Collecting field data

The crop cover type of every sample pixel was obtained through fieldwork conducted in middle-to-late August 2015. This time window was chosen because soybean plants in the U.S. typically reach reproductive stages with maximum canopy cover during this time of year. This optimal time was confirmed by satellite data. Soybean pixels exhibit maximum normalized difference vegetation index (NDVI) (Tucker, 1979) on Julian date 225 in the MODIS 16-day composites, corresponding to August 13–28 (Fig. 2).

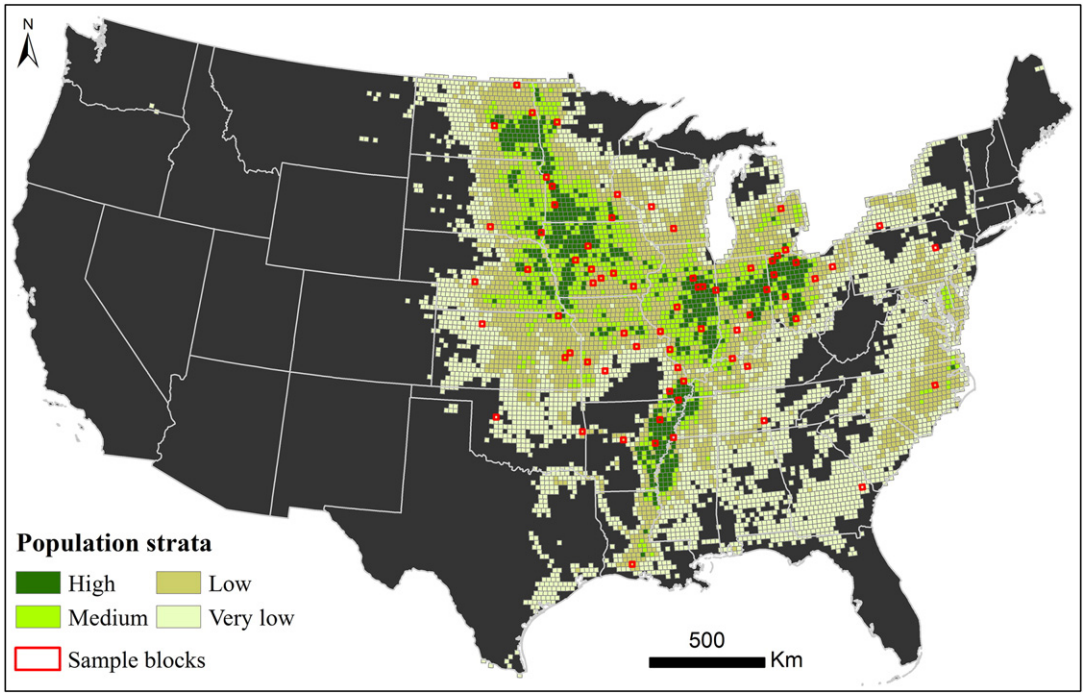

Fig. 1. Population strata and sampled blocks for soybean area estimation in the United States.

Because of the difficulty of visiting sample pixels located within large privately-owned properties in the Midwest U.S., field boundaries containing the sample pixels were manually drawn prior to fieldwork. A field polygon was delineated based on spectrally homogeneous pixels in early August, an indicator of a single crop within the field (Fig. 3). Access to each field was marked, typically where roads coincided with a field boundary.

We visited all sample pixels in the field and recorded information on the observation coordinates, the crop type, coverage, height, growing stage as well as any observable environmental conditions such as poor soil or signs of flooding history. We also took geo-tagged photos at each field site. The collected data were double-checked for quality assurance after fieldwork, focusing on three aspects of the data collection process. First, field boundaries could have been imprecisely drawn due to cloud or shadow presence or Landsat 7 Enhanced Thematic Mapper Plus (ETM+) scan line corrector error (SLC-off) effect. These boundaries were re-edited when clear-view OLI images became available in the Landsat database. Second, 31 out of the 700 sample points were inaccessible in the field. The crop type (soybean vs. non-soybean) of each inaccessible point was inferred by interpreting time-series of Landsat spectral reflectance as well as using investigators' local knowledge obtained through fieldwork. Third, 15 out of the 700 sample points were located on the boundary between soybean and non-soybean fields and contained mixture of crop types. We estimated the proportion of soybean cover within each mixed 30 m × 30 m pixel by overlaying the Landsat pixel footprint on high-resolution images in Google Earth.

Table 1
Stratum definition and sample size for U.S. soybean area estimation. Note: national soybean coverage is derived from 2010–2014 CDL and block size is 20 km × 20 km.

Stratum	National soybean coverage	Soybean intensity per block	Number of blocks in stratum	Number of blocks sampled
High	33%	>31.7%	698	20
Medium	33%	20.32–31.7%	1002	20
Low	29%	4.2–20.32%	2111	20
Very low	4.9%	<4.2%	3217	10
Total	99.9%		7028	70

2.1.3. Estimating soybean area from the field sample

The soybean area and associated uncertainty were estimated using stratified two-stage cluster sampling formulas. Soybean area estimates were produced for each of the four strata, and these stratum totals were then summed to obtain the U.S. total estimate. The details of the estimation protocol are described as follows. Define $A_{h,i}$ to be the area of cropland in block (cluster) i of stratum h, and define $\hat{p}_{h,i}$ to be the estimated proportion of soybean from the sample points within the cropland area of block i of stratum h. The area of soybean in each block ($\hat{Y}_{h,i}$) is estimated as

$$\hat{Y}_{h,i} = A_{h,i}\hat{p}_{h,i}. \quad (1)$$

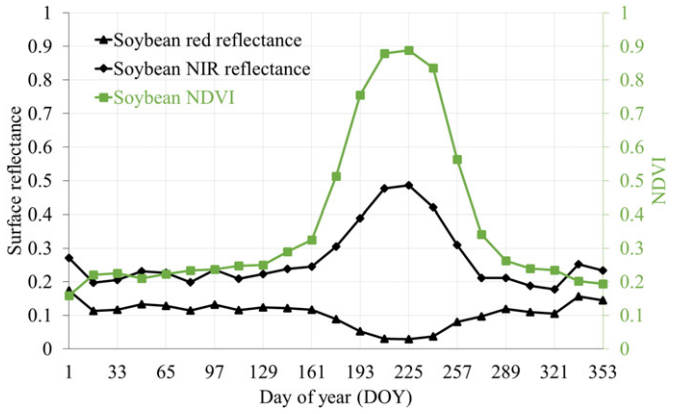

Fig. 2. MODIS-based time-series of soybean NDVI, red and near-infrared (NIR) reflectance in the United Sates. Each point represents the median value of all pure soybean pixels at the 250 m resolution. Data are from CDL and 16-day MODIS surface reflectance composites of year 2013.

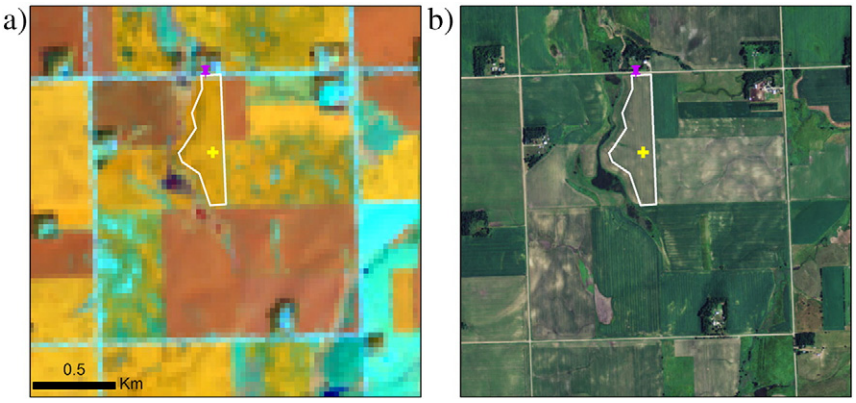

Fig. 3. Delineating field boundaries using Landsat and high-resolution images. a) Landsat 8 Operational Land Imager (OLI) data acquired on August 4, 2015, showing in R: B5, G: B6 and B: B4 combination. b) High-resolution image in Google Earth. The yellow cross on both a) and b) indicates the location of a sample pixel. The white polygon delineates the boundary of a homogeneous field drawn by an investigator prior to the field visit. The magenta pin marks the location where investigators made the actual field observation. The center coordinate of this example is (95.9269° W, 44.2181° N).

The total area of soybean in stratum h ($\hat{Y}_h$) is estimated by combining the block area estimates using

$$\hat{Y}_h = \frac{N_h}{n_h}\sum_{i=1}^{n_h}\hat{Y}_{h,i}, \tag{2}$$

where n_h and N_h are the sample and population sizes for stratum h. Lastly, the total soybean area ($\hat{Y}$) is estimated by summing the area estimates over all four strata:

$$\hat{Y} = \sum_{h=1}^{4}\hat{Y}_h. \tag{3}$$

The estimated variance for the stratum-specific soybean area estimate is

$$\hat{V}\left(\hat{Y}_h\right) = N_h^2\left(1-\frac{n_h}{N_h}\right)s_h^2/n_h, \tag{4}$$

where s_h^2 is the sample variance of the n_h values of $\hat{Y}_{h,i}$ in stratum h. The estimated variance for total area of soybean estimated for the U.S. is the sum of the four stratum-specific variance estimates:

$$\hat{V}\left(\hat{Y}\right) = \sum_{h=1}^{4}\hat{V}\left(\hat{Y}_h\right). \tag{5}$$

The standard error (SE) of an area estimate is the square root of the estimated variance. As is typically the case in practice, the second-stage (i.e., within block) contribution to variance was not estimated because it is negligible relative to the first-stage variance (Lohr, 2010, p. 185).

We estimated the soybean area and associated standard error using three groups of data: (1) field-verified pure pixels ($n = 654$), (2) field-verified pure pixels and image-interpreted pixels ($n = 685$) and (3) field-verified pure pixels, image-interpreted pixels and mixed pixels ($n = 700$). Excluding inaccessible and mixed pixels affected the estimates of $\hat{p}_{h,i}$ and thus the total area estimates. However, as shown later, excluding inaccessible and mixed pixels did not substantially change the estimated total area or the standard errors.

2.2. Landsat-based wall-to-wall soybean mapping

2.2.1. Data processing

All Landsat images covering the study area (i.e. the block population in Section 2.1.1) and acquired during the summer growing season defined as from June 1st to September 15th were employed for soybean classification. A standardized Landsat data processing system has been developed, tested and applied previously in a number of applications at national to global scales (Hansen et al., 2013; Hansen et al., 2008a; Potapov et al., 2014; Potapov et al., 2012; Potapov et al., 2015). Each Landsat image is per pixel quality assessed and radiometrically normalized; the stack of viable growing season observations is used to derive a set of phenological metrics, the inputs to the mapping algorithm. Details are described as follows.

A total of 3409 Landsat 7 ETM+ and Landsat 8 OLI and Thermal Infrared Sensor (TIRS) L1T images were downloaded from the United States Geological Survey Earth Resources Observation and Science (USGS EROS) data center. The images were converted from digital numbers to top-of-atmosphere (TOA) reflectance and brightness temperature using the standard protocol reported in Chander et al. (2009). The TOA reflectance images were screened for masking cloud, shadow, haze and water using an array of quality assessment models, which were essentially a set of pre-defined, generic bagged decision tree models. The quality-flagged TOA images were normalized using MODIS surface reflectance for correcting the atmospheric effect as well as the Landsat across-track anisotropy effect (Potapov et al., 2012). To correct the atmospheric effect, we calculated a mean bias between MODIS surface reflectance and Landsat TOA reflectance for each spectral band over the image area consisting of good quality land observations, and applied the bias value to adjust Landsat reflectance. To correct the across-track viewing anisotropy effect, we regressed the Landsat scan angle against the bias between Landsat and MODIS and applied the regression coefficients to further adjust Landsat reflectance. This radiometric normalization step was conducted for red, near-infrared and shortwave infrared (SWIR) reflective bands. Thermal bands (ETM+ band 6–2 and TIRS band 10) were converted to brightness temperature with no additional normalization. Other visible bands were excluded due to residual atmospheric contamination, most notably strong scattering of visible radiation by the atmosphere. In addition to these spectral bands, we also calculated NDVI and normalized difference water index (NDWI) (Gao, 1996) for every Landsat pixel.

A set of multi-temporal metrics were generated using the radiometrically normalized images. Multi-temporal metrics are statistical transformations of image time-series. They can capture the salient features of vegetation phenology while maintaining high spatial and temporal data consistency. Hence, they are suitable for land cover characterization at continental to global scales (DeFries et al., 1995; Hansen et al., 2000). Here we provide the first test of seasonal Landsat multi-temporal metrics (Fig. 4) for national scale, wall-to-wall crop-type mapping. We generated two groups of metrics: (1) rank-based metrics representing selected percentiles (the minimum, 10%, 25%, 50%, 75%, 90% and the maximum) of each reflective band as well as reflectance values corresponding to selected percentiles of ranked NDVI, NDWI and brightness temperature; (2) average metrics that were calculated as the mean

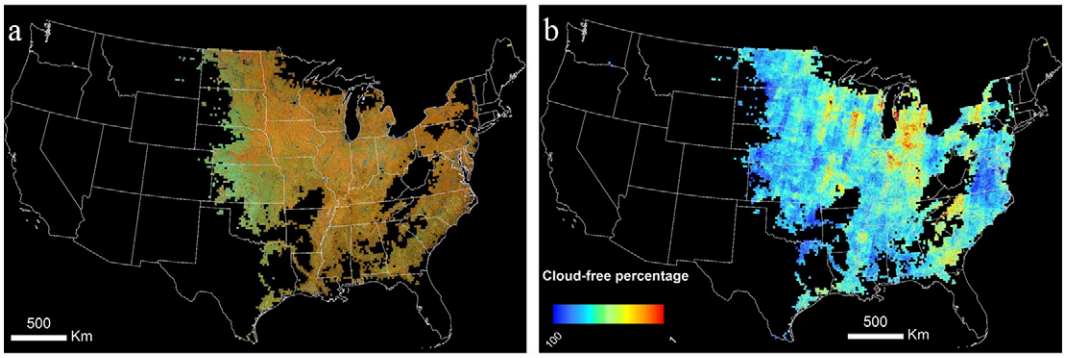

Fig. 4. Landsat data over summer-crop growing season. (a) Image composites with RGB channels representing 90-percentile NIR reflectance, 10-percentile SWIR1 reflectance (band 5 of ETM + or band 6 of OLI) and 10-percentile red reflectance, respectively. (b) Percentage of cloud-free observations. White lines delineate state boundaries.

value of all observations between two selected percentiles for each reflective band as well as the mean reflectance value of all observations located between two selected percentiles of ranked NDVI, NDWI and brightness temperature. These metrics were designed to represent key characteristics of vegetation phenology, which altogether made up a comprehensive feature space.

2.2.2. Soybean classification

U.S. soybean was mapped using a supervised bagged classification tree model (Breiman et al., 1984). Training data were derived through visual image interpretation. Because of the high diversity of crops as well as varying growing conditions of each crop across the U.S., it is crucial to ensure that the training dataset captures a wide range of geographic and ecological variability. Five types of information were interactively displayed to aid image interpretation and training selection: (1) Landsat image composites corresponding to the peak of the growing season, (2) high-resolution images from Google Earth, (3) MODIS 8-day NDVI time-series, (4) historical CDL crop type maps and (5) 2015 field sample. Each training pixel was labeled as either soybean or non-soybean. The extremely high NIR reflectance of soybean at the peak of the growing season is an important feature for differentiating soybean and corn in the "corn belt" region where corn-soybean mosaic is the most common crop mixture (e.g. Iowa, Illinois, Indiana and Ohio). Precise training polygons were drawn by cross-checking field boundaries on high-resolution images. Historical CDL crop maps and 2015 field sample jointly provided information about the diversity of crops in a particular region, which is critical for ensuring the representativeness of the training set. In addition, crops that have similar spectral response as soybean at the peak of the growing season may have a different growing window or different phenological transitions, which can be visualized using MODIS time-series. For example, dense alfalfa in mid-August may have similar NIR reflectance as soybean but this perennial crop has a longer growing season. Over two million pixels drawn as polygons were iteratively selected for training the classification tree.

Classification tree models have been widely used in land cover mapping (Friedl and Brodley, 1997; Hansen et al., 1996). This non-parametric machine learning method is chosen because it is flexible for handling nonlinear relationships between class membership and input feature and the decision rules are intuitive for human interpretation. At the training stage, the tree grows by recursively splitting training data into less heterogeneous partitions until criteria of accuracy or purity are met. The model is fully constructed when all leaf nodes are generated. In the case of binary classification, each leaf node of the tree consists of either one pure class or a mixture of two classes, the probability values of which are determined by the relative proportions of training pixel counts in the leaf node. Driven by training data, tree models tend to overfit. Thus, a bagging procedure was applied to improve model stability and prediction accuracy. We trained a bagged classification tree model (7 trees) using the binary soybean and non-soybean training dataset and predicted soybean probability (median of the 7 outputs) per Landsat pixel across the study area.

A novel feature introduced into the application of the classification tree for cropland mapping in this study was to search and apply optimal probability thresholds to convert the probability layer to a binary soybean and non-soybean classification map such that mapped soybean area matched the sample-based area estimate. Traditionally, binary classification using a decision tree model applies a 0.5 probability threshold to determine the predicted class membership. Discretizing the probability layer with flexible thresholds can achieve a number of objectives such as matching a sample-based area estimate, balancing user's and producer's accuracy of a target class, or matching the prior probability of a class distribution (Broich et al., 2011; Bwangoy et al., 2010). Here the map calibration step was conducted for each stratum (Fig. 1) (i.e. a probability threshold was determined for each stratum). Soybean area in a stratum derived through pixel counting was computed as a function of the probability threshold; a higher threshold yielded a lower area estimate and vice versa. By progressively slicing the threshold from 0 to 1, a threshold value was chosen such that the desired mapped area of soybean matched a specified soybean area (e.g., the sample estimated area). A national soybean classification map was then produced by applying the four stratum-specific thresholds to each pixel in its respective stratum. For comparison, a soybean map resulting from a 0.5 probability threshold was also produced.

2.2.3. Soybean map validation

The field sample was used as reference data to validate soybean mapping results. Overall accuracy and per class user's accuracy and producer's accuracy were estimated and the weight of each sample pixel was the inverse of its inclusion probability (Stehman et al., 2003). The formulas for estimating accuracy are presented in the Appendix. Confusion matrices were estimated for both the original map created using the 0.5 probability threshold to determine soybean and for the calibrated map. While the confusion matrix for the map based on the 0.5 probability threshold represents a completely independent validation, the confusion matrix for the calibrated map is likely somewhat optimistic because the probability threshold used to produce this map was based on the reference data. However, we wanted to examine how the error structure of the calibrated map changed as a result of the calibration.

3. Results

3.1. Sample-based national soybean area estimates

Soybean cultivation area estimated from the field sample was 341,000 km^2 with a SE of 23,000 km^2 (Table 1). This result is 1.0%

Table 2
Soybean cultivation area (km^2) and uncertainty estimates based on field sample data obtained for a probability sampling design.

Stratum	Pure pixels (n = 654)		Pure pixels and image-interpreted pixels (n = 685)		Pure pixels, image-interpreted pixels and mixed pixels (n = 700)	
	Area	SE	Area	SE	Area	SE
High	98,273	9844	98,975	9658	99,598	9442
Medium	96,690	8960	94,298	8611	94,783	8086
Low	117,903	17,669	120,737	16,253	120,738	16,331
Very low	25,833	11,348	25,640	11,380	25,797	11,372
Total	338,699	24,862	339,649	23,687	340,916	23,463

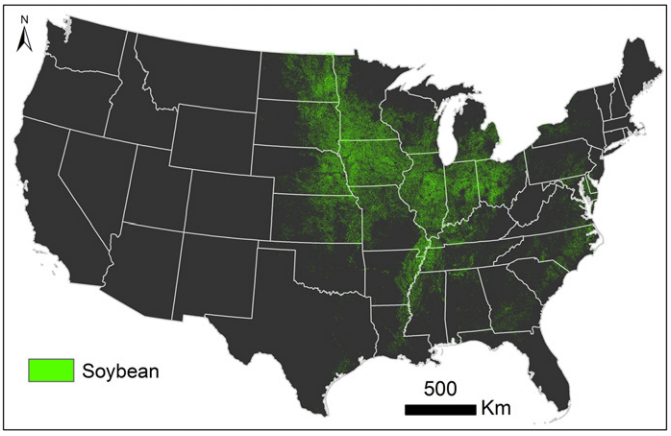

Fig. 5. Soybean classification map over the United States in year 2015.

lower than the area of planted soybeans reported by USDA NASS's 2015 June enumerative survey (344,000 km^2) and 1.9% higher than NASS's 2016 January estimate (334,000 km^2) (http://www.nass.usda.gov/) although both NASS values are within one standard error (SE) of our field-based estimated soybean area. The high and medium strata accounted for 29% and 28% of the estimated total soybean area, respectively, both with small SE. The largest per-stratum SE was found in the low stratum, which accounted for 30% of the study population (Fig. 1) and 35% of the total estimated soybean area. The very low stratum covered 46% of the study region but accounted for only 7% of the total estimated soybean area. Its contribution to the total SE was also comparatively low. We also found that excluding image-interpreted and mixed pixels did not substantially change the area or the SE estimates (Table 2).

3.2. Soybean classification map and accuracy

The calibrated soybean map of the United States in year 2015 is shown in Fig. 5. At the national scale, the known spatial patterns of soybean cultivation are evident—as the concentrated distribution of soybean is readily identifiable in the Midwest states, the Great Plains states, the lower Mississippi valley and the east coast. Fig. 6 shows three representative sites in Iowa, Arkansas and North Carolina, which depict more spatial details at the field scale. The agricultural landscape in the Midwest states such as Iowa typically consists of rectangular mosaics of corn and soybean fields (Fig. 6d) and the pixel-based classification yields homogeneous fields with clear boundaries (Fig. 6a). The Mississippi delta region shows a greater crop diversity, smaller field sizes and a variety of environmental conditions (Fig. 6e), which result in some salt-and-pepper effect noticed on the map (Fig. 6b).

The soybean area derived from the uncalibrated map through pixel counting was 260,000 km^2, 24% lower than the sample-based estimate, whereas the area estimate derived from the calibrated map was 341,000 km^2, exactly matching by construction the sample-based estimate. An overall accuracy of 84% was achieved for the uncalibrated map (Table 3). However, both producer's accuracy (53%) and user's accuracy (77%) of the soybean class were markedly lower than the non-soybean class (producer's accuracy: 95%, user's accuracy: 86%). The calibration step not only changed the total soybean area estimate but also

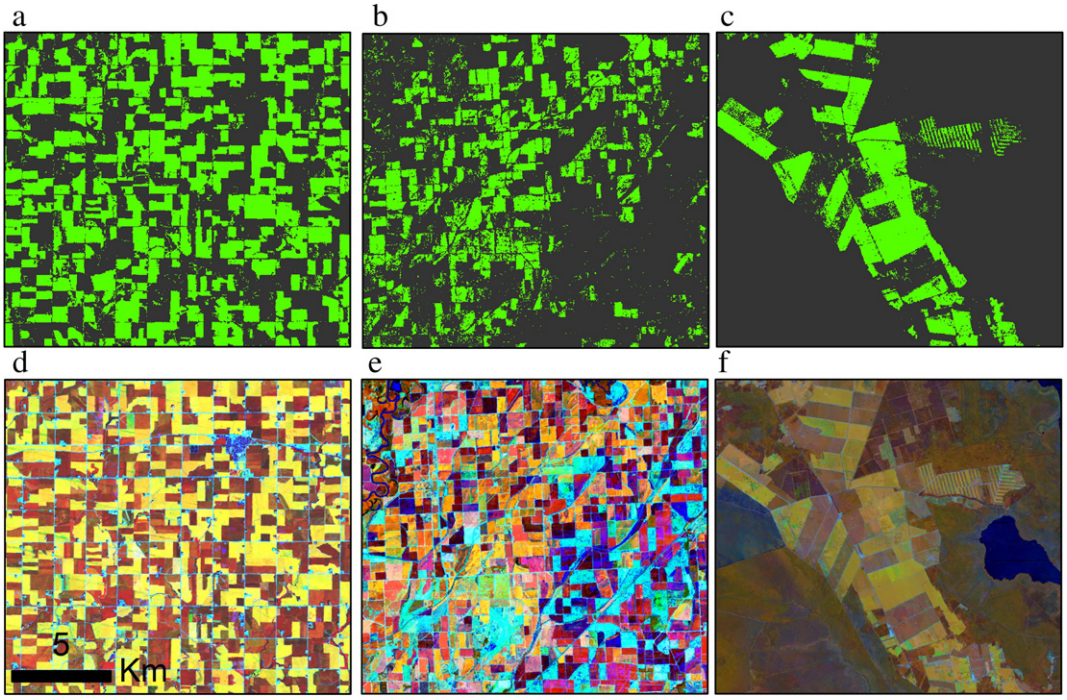

Fig. 6. Soybean classification in selected representative sites in the United States. Panels (a), (b) and (c) are classifications in Iowa, Arkansas and North Carolina and panels (d), (e) and (f) are Landsat image composites (R: NIR, G: SWIR1, B: Red) of the three sites at the peak of the growing season. All panels are displayed at the same scale (18 km × 18 km). The center coordinates of (a), (b) and (c) are (95.841° W, 42.783° N), (90.907° W, 36.169° N) and (76.164° W, 35.805° N), respectively.

Table 3
Confusion matrices of the soybean classification maps (cell entries represent proportion of area). Reference data were obtained from a probability sample and the cover type of each sample pixel was determined by field survey. The uncalibrated map was generated by applying a 0.5 threshold to the soybean probability layers produced by the decision tree model, whereas the calibrated map was generated by applying stratum-specific thresholds that matched mapped soybean area with sample-based soybean area estimates.

	Class	Reference			User's accuracy % (SE)	Producer's accuracy % (SE)	Overall accuracy% (SE)
		Soybean	Non-soybean	Sample total			
Uncalibrated map	Soybean	0.13	0.04	0.17	77 (5)	53 (4)	84 (2)
	Non-soybean	0.12	0.71	0.83	86 (2)	95 (1)	
	Map total	0.25	0.75	1.00			
Calibrated map	Soybean	0.16	0.05	0.22	75 (5)	64 (4)	86 (2)
	Non-soybean	0.09	0.69	0.78	88 (2)	93 (2)	
	Map total	0.25	0.75	1.00			

altered the error structure of the map, especially for the soybean class. Producer's accuracy of the soybean class increased from 53% to 64% after calibration with a small penalty on the user's accuracy, which declined from 77% to 75%. The overall accuracy also slightly increased from 84% to 86% after calibration. Since the sample-based area estimates were used in making the calibrated map, the accuracy assessment of the calibrated map does not represent the accuracy derived from another independent sample. However, the increase in overall and producer's accuracies and the decrease in user's accuracy of the soybean class were likely true effects of the calibration.

4. Discussion

4.1. The efficiency of probability sampling for crop area estimation

We described two different approaches for deriving crop-specific information in a country. The sampling method has the advantage of producing unbiased area estimates in an efficient and transparent manner whereas wall-to-wall mapping provides spatially explicit crop cover information. The interactions of the two methods are: (1) historical wall-to-wall crop maps are used to construct strata for sample selection, (2) the independent field sample can be directly used for conducting a statistically rigorous validation, and (3) a classification map can be adjusted so that the mapped area matches the sample-based area estimate. As a result of the synergistic use of sample field data and satellite imagery, the disparity between sample-based area estimates and the mapped area, which is common in current land cover and land use change research, is substantially reduced.

Our two-stage cluster sampling estimate was only 1.9% higher than the NASS estimate indicating that the protocol of stratification and field sampling could produce an estimate nearly two months ahead of harvest that was reasonably close to the NASS estimate. The sample of 70 clusters (blocks) is a relatively small sample size and so the standard error for estimated total soybean area was 6.9% of the estimated total. In an operational setting, a larger sample size of clusters would be needed to reduce the standard error to a more tolerable level. A larger sample size per stratum would also allow use of model-assisted estimators to reduce the standard error of the sample-based area estimators (Gallego, 2004; Stehman, 2009). An additional 25–30% reduction in standard error could be expected by applying regression estimators (Gallego et al., 2014; Gonzalez-Alonso et al., 1997; King et al., in review). Because our maximum sample size for a stratum was only 20, we did not evaluate any model-assisted estimators to avoid the possibility of incurring bias (Cochran, 1977, Sec. 7.11).

The stratification implemented using CDL was effective as demonstrated by the fact that a simple random sample of 70 clusters (i.e. without CDL for stratification) would have yielded a SE that was 1.92 times larger than the SE of the stratified design. We also conducted a retrospective assessment of our choice of sample allocation to strata. Using the estimated per-stratum standard deviations from the sample data, the optimal allocation of sample size to strata to minimize the variance of the estimated U.S. total soybean area would have been 16, 14, 27 and 13 for the high, medium, low and very low strata. In reality this optimal allocation was not very far off from the 20, 20, 20, and 10 sample size allocation that was used.

The efficiency of the sample-based area estimation stems from the spatially explicit stratification. Here, remote sensing-derived near-term historical soybean maps from NASS were used to target field visits for national-scale area estimation. Our field data collection was completed by 11 investigators with 4 vehicles in <2 weeks and resulted in an estimate very close to that of NASS by early September. Implementing this sample-based framework in other countries or regions would require such spatial data to exist or to be created otherwise. The CDL program, although producing highly accurate crop maps, may be difficult to operate in other countries due to its dependency on comprehensive census-based data for training. While national-scale mapping for crop area estimation remains challenging, we have shown that the classification procedure using standardized time-series metrics and expert-interpreted training could generate a soybean map with reasonable accuracy (84–86% overall accuracy). Such a procedure can be easily applied in other regions to produce historical or current crop maps for the purpose of stratification in a future year. Thus our approach shows great potential for application in developing countries where advanced mapping and survey systems do not exist or are too expensive to be established.

4.2. Decision tree model, map calibration and error analysis

The branch structure of one of the seven bagged classification tree models is illustrated in Fig. 7. The root split utilizes the most informative metrics mean90max_NIR_NDVI (mean of NIR reflectance of all observations between 90-percentile and maximum NDVI) to separate summer/spring/perennial herbaceous crops such as soybean and cotton from non-crop land covers such as natural vegetation, wetland, barren land, water, as well as winter crops and tall summer crops such as corn and sorghum. Following the splitting rules down the path, the model represents a hierarchical procedure for crop type classification, although the specific type of non-soybean crop is not explicitly identified in the training dataset. The left-child branch discriminates darker surfaces (e.g. water, wetland, forests) from brighter land surfaces (e.g. urban, herbaceous vegetation, shrubs) using peak SWIR2 reflectance (mean75max_SWIR2). Maximum NDVI is then used to distinguish tall summer crops (e.g. corn, sorghum) from natural herbaceous vegetation, winter/spring crops as well as non-vegetated land. In the right-child branch, broadleaf summer crops (e.g. soybean, cotton) are first distinguished from perennial crops (e.g. alfalfa) using the metrics mean75max_SWIR2_NDVI indicating plant water content during peak greenness. Spring crops such as spring wheat and some summer crops (e.g. sugarbeets) in high-latitude drained wetlands in North Dakota and Minnesota, which exhibit extremely low SWIR2 reflectance at peak greenness, are also grouped with perennial crops. Then, further down the next right-child branch, the metrics max_red_NDWI (red reflectance corresponding to max NDWI) indicating crop photosynthetic activity at peak growth stages (e.g. soybean R3–R5) are used for further split.

Note that only major branches of the tree model are displayed and none of the nodes illustrated is a leaf node in Fig. 7. Soybean pixels

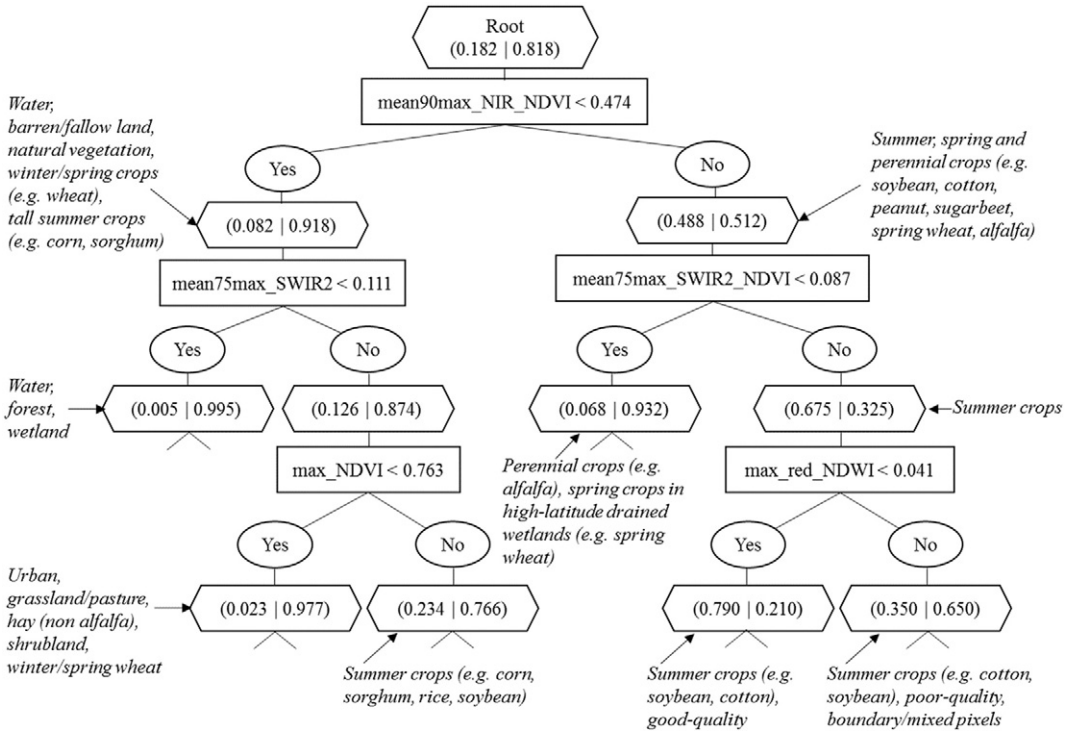

Fig. 7. Decision tree model for soybean classification in the United States using time-series metrics. Hexagon: tree nodes (soybean proportion | non-soybean proportion), rectangular: decision rules. This tree has a maximum depth of 21 and a total of 471 leaf nodes. Only major branches are shown in this figure. mean90max_NIR_NDVI: mean of NIR reflectance of all observations between 90% and maximum NDVI; mean75max_SWIR2: mean of 75% and maximum SWIR2 reflectance; mean75max_SWIR2_NDVI: mean of SWIR2 reflectance of all observations between 75% and maximum NDVI; max_NDVI: maximum NDVI value; max_red_NDWI: red reflectance value corresponding to maximum NDWI.

could be mixed in every intermediate node due to a variety of geographic and ecological factors related to the spectral properties of soybean cover. These diversities were addressed by other metrics further down the expanding paths. A total of 471 leaf nodes were eventually grown in this 21-layer tree. Vegetation indices themselves are useful for separating general land covers. For crop type mapping, vegetation indices are more useful for providing ranking mechanisms partly because they are more robust to band-correlated noise (Huete et al., 2002); the corresponding spectral reflectance values are then used to do the split. These results suggest that time-series metrics represent well the fundamental biophysical properties of vegetation, which are closely related to specific crop phenology and crop type. Because the utility of these metrics is not limited to specific locations or time, the metrics are also useful for generalizing tree-based classification analysis to other regions of the globe.

The immediate output of the tree model is a per-pixel soybean probability layer (Fig. 8). Soybean area derived from a classification map is clearly a function of the probability threshold with area decreasing monotonically against an increasing threshold. The probability thresholds that resulted in the mapped area matching the sample estimated area (by stratum) were 0.20, 0.35, 0.19 and 0.47 for the high, medium, low and very low stratum.

We also translated the uncertainty of the sample estimate to a map leveraging the probability layer. For the purpose of illustration, we chose the probability thresholds corresponding to the 85% confidence interval for soybean area estimated from the sample (estimate $\pm$ 1.44 · SE). That is, the stratum-specific thresholds to produce the soybean map were chosen so that the area mapped as soybean matched the lower and upper 85% confidence bounds of the sample-based area estimate. The thresholds that resulted in the lower bound soybean area were 0.32, 0.47, 0.29 and 0.83 and the thresholds that resulted in the upper bound soybean area were 0.11, 0.25, 0.12 and 0.26 for the high, medium, low and very low stratum, respectively. Large homogeneous soybean fields were characterized with high soybean probability whereas some small patches within large fields and isolated speckles were characterized with low soybean probability (Fig. 9). We also

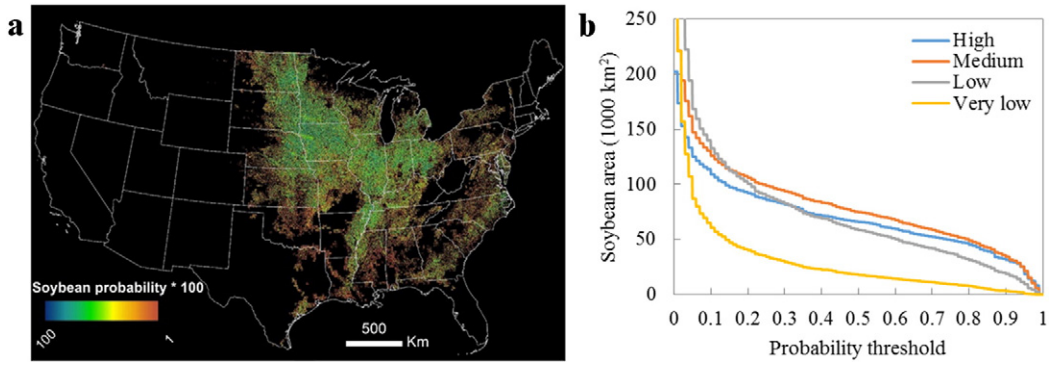

Fig. 8. (a) Soybean probability layer output by the classification tree model. (b) Map-based soybean area as a function of probability threshold.

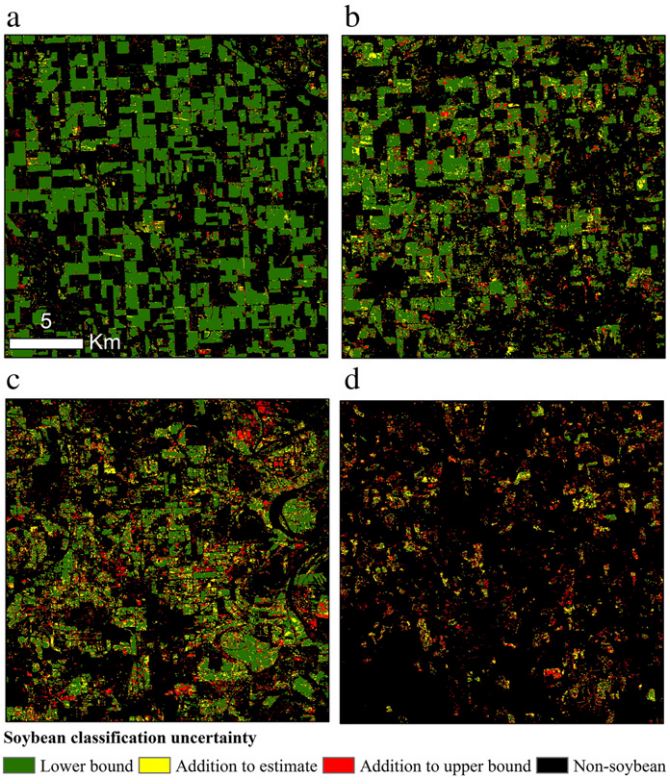

Fig. 9. Soybean classification uncertainty at 30 m pixel level at selected representative sites. Green: soybean pixels with high probability that make up the lower bound area; Yellow: additional soybean pixels that match the sample estimate; Red: additional soybean pixels with low probability that match the upper bound area. a) central Iowa; b) western Illinois; c) western Mississippi; d) southern Alabama. All panels are displayed at the same scale. The center coordinates of a), b), c), and d) are (92.692° W, 42.941° N), (90.426° W, 40.597° N), (90.402° W, 33.757° N) and (86.103° W, 31.235° N), respectively.

derived the lower bound soybean area, the upper bound area, sample estimated area, and computed the ratio of uncertainty range (upper bound – lower bound) and sample estimate for each 20 km × 20 km block (Fig. 10). Note the close resemblance of this uncertainty map and the stratification map (Fig. 1), where the high, medium and low strata had relatively low uncertainty but the very low stratum had relatively high uncertainty. These results highlighted the challenge of mapping soybean in dry environments (e.g. Kansas, Oklahoma and Texas) and in small fields in the Appalachian forested landscapes (e.g. Pennsylvania, Tennessee and Alabama).

A number of factors could have contributed to the soybean omission and commission errors. Poor quality data as a result of cloud and shadow contamination could considerably change the reflectance value, especially when data were missing at the critical periods of time for classification (e.g. peak of the growing season). Although agriculture fields in the U.S. are large (Yan and Roy, 2016) and most pixels contain only one crop type, edge pixels, where multiple crops, grass, trees, road and water could be mixed, may have lower accuracy. The low quality of some soybean fields was a major contributor to omission errors. Ground observations suggested that at the time of the fieldwork (mid-to-late August), some soybeans in the Mississippi delta, which were omitted by the map, were at the flowering stage instead of the reproductive stage as most other soybeans were in the U.S. Plants could also fail due to environmental stresses such as drought, flood and other abnormal weather events. These error factors imply that implementing the developed mapping approach in other agricultural regions would encounter similar challenges, especially in crop diverse regions comparable to North Dakota or the Mississippi delta. In particular, tropical agriculture typically has a great diversity of crop types, different crop cycles (normal, precocious, super-precocious), different agricultural calendars and a variety of management techniques (e.g. no-tillage, irrigation). One option to address such complexity is to stratify the study area in order to account for landscape-specific conditions. It has been demonstrated that locally-tuned models often employ different spectral information than national-scale models, improving mapping accuracy (King et al., in review).

An estimated 36% of commission errors in the calibrated map were from corn, 27% from grassland/pasture, 20% from alfalfa, 7% from cotton, 3% from fallow land, 3% from bare land, and the rest from sorghum, wheat and urban/suburban land covers, each contributing <1% (these contributions incorporate the estimation weights of each sample pixel for the stratified design). Compared with correctly classified soybean pixels, both omission errors and commission errors show marked differences for key metrics (Fig. 11). For mean90max_NIR_NDVI, the most

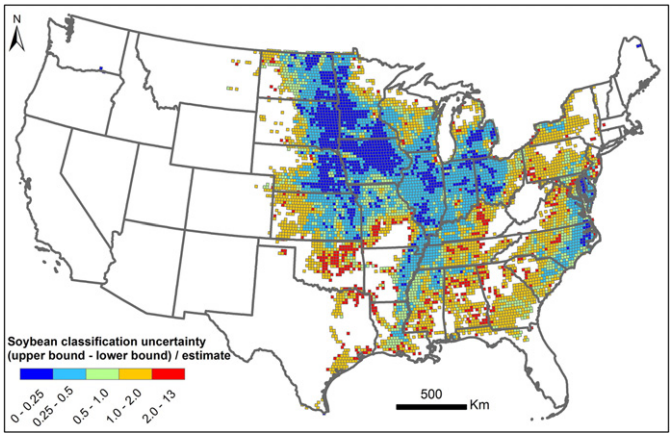

Fig. 10. Soybean classification uncertainty at 20 km block level. For each block, the range of uncertainty (upper bound soybean area – lower bound soybean area) was normalized by the sample estimated area. Thus, a value close to 0 indicates relatively low uncertainty and a larger value indicates relatively high uncertainty (i.e., high uncertainty indicates a block where a large change in the area of soybean occurs as the probability threshold is varied).

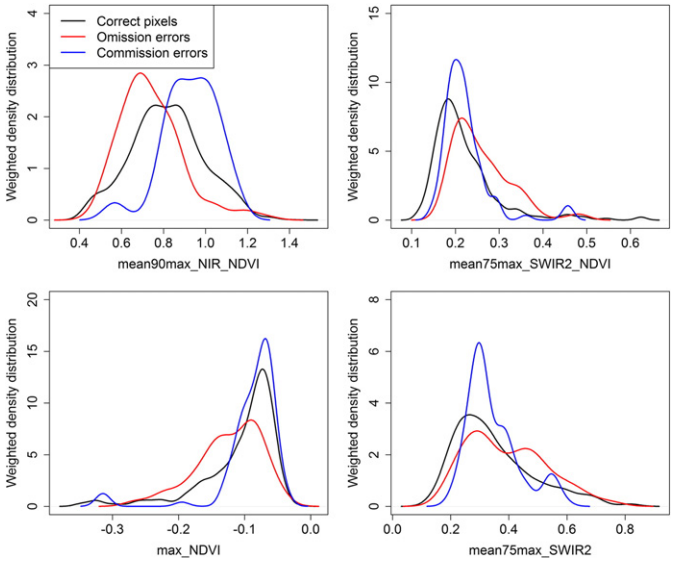

Fig. 11. Weighted density distributions of the four most heavily used metrics by the classification tree model for validation sample of correctly classified pixels (black), omission errors (red) and commission errors (blue). The area under the curve sums to 1 for all curves in the plot.

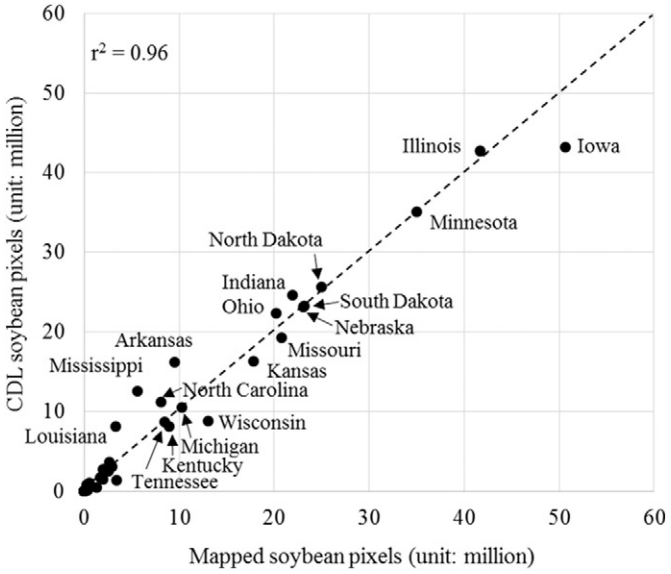

Fig. 12. Comparing soybean map generated in this study and CDL at state level.

heavily used metrics by the decision tree, omitted soybean pixels show significantly lower values than correct pixels whereas commission errors show significantly higher values.

4.3. Map comparison with the cropland data layer

In addition to the validation exercise, which reveals the overall error structure of the derived soybean map, here we provide a detailed comparison with the 2015 CDL to show spatial patterns of error and to further illustrate the strength and weakness of our classification methodology. The total area of all soybean pixels in our calibrated classification map is 341,000 km^2, 0.6% lower than the 343,000 km^2 derived from CDL. State-level comparison of soybean area has an r^2 of 0.96 and a mean absolute difference of 1127 km^2 with most states located close to the 1:1 line (Fig. 12). The largest differences occur in Iowa and Wisconsin in which our map identifies more soybean, as well as in Arkansas, Mississippi and Louisiana in which our map identifies less soybean. Comparison is also conducted at the block scale (Fig. 13) based on the ratio of the difference of soybean percentage divided by CDL soybean percentage per block (positive values indicate more soybean in our map whereas negative values indicate more soybean in CDL). Clustered regions of disagreement are clearly shown, notably Wisconsin, Georgia, Arkansas and Mississippi. Pixel-level investigation suggests that the major crop causing overestimation of our map is alfalfa in Wisconsin and cotton in Georgia. The large difference in the Mississippi delta is due to classification errors in both datasets (based on visual examination of Landsat) (Fig. 14). Both maps identify and agree on most soybean fields regardless of single or double cropping but differences appear along field edges and in certain fields with distinct spectral signatures. Our map shows some soybean pixels in CDL-labeled sorghum fields (image center), whereas CDL shows soybean pixels in fields with low vegetation cover (lower-left part of the image). Future research would be focused on improving soybean mapping in difficult regions as well as extending the developed method for classifying multiple crops.

5. Conclusions

This paper demonstrated a method for estimating cultivation area for a specific crop as well as producing spatially explicit crop cover at

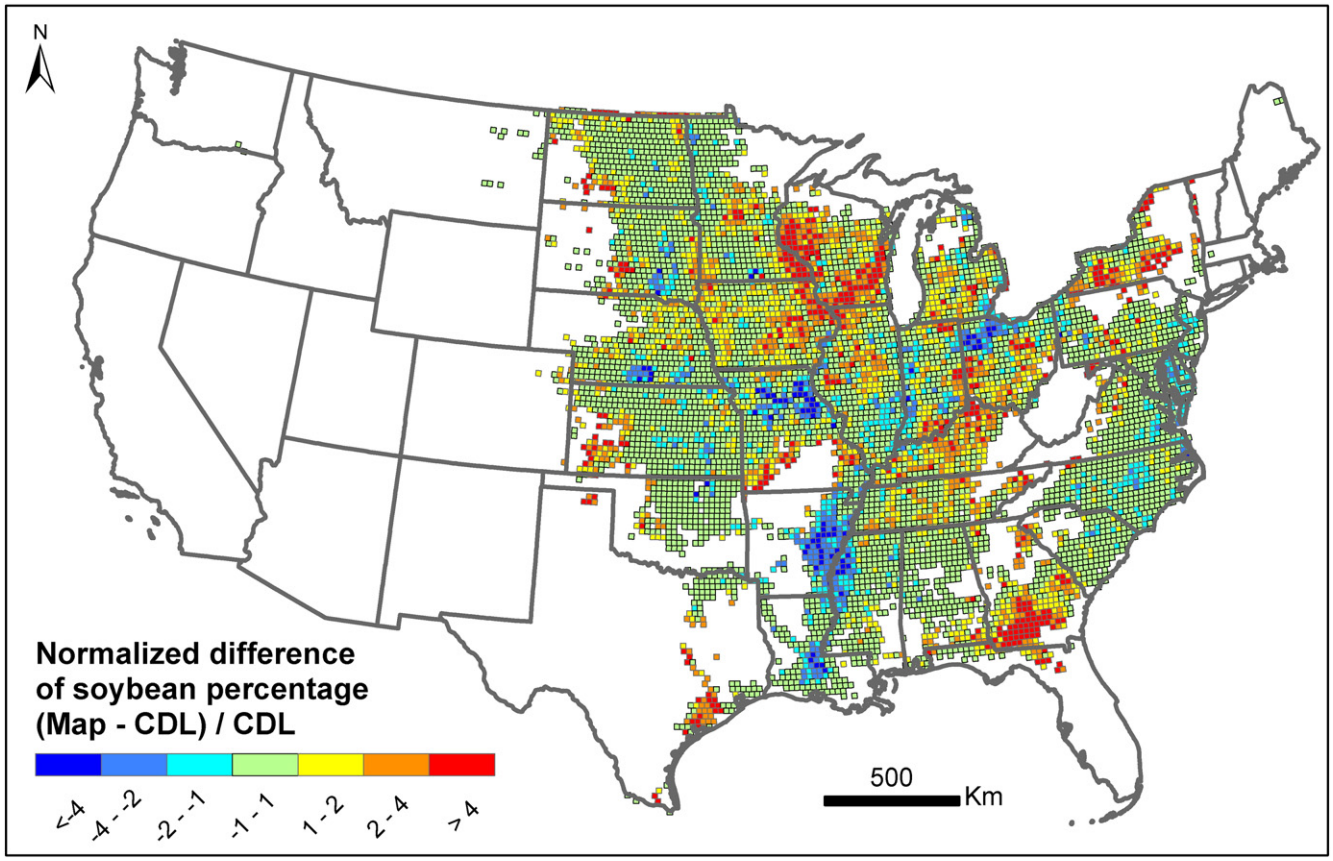

Fig. 13. Comparing soybean map generated in this study and CDL at block level.

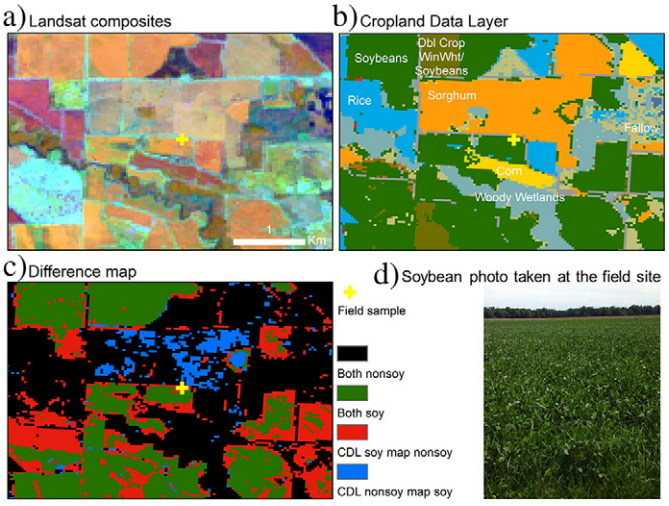

Fig. 14. Comparing soybean map of this study and CDL at pixel level at a field site in the Mississippi delta region. a) Landsat composites with RGB channels representing NIR, SWIR1 and red reflectance at the peak of the growing season. Various colors indicate the crop diversity in this region. b) USDA CDL in year 2015. c) Difference map. Agreement pixels are shown in black (non-soybean) and green (soybean) whereas disagreement pixels are shown in red and blue. d) Soybean photo taken at the field site marked by a yellow cross in a), b) and c). The coordinate of this field sample is (91.0728° W, 34.8450° N) in eastern Arkansas.

national scales. The method consists of two interdependent parts: (1) a probability sampling design for field data collection and area estimation and (2) wall-to-wall Landsat data processing for crop mapping. While the two approaches were implemented separately, the field-based sample estimated area was used to calibrate the map product to align the soybean area of the map with the sample estimated area. Results of the study include a U.S. 2015 soybean cultivation area estimate with associated estimated standard error and a map characterization of soybean cover with associated spatial uncertainty.

Our stratified two-stage cluster sampling design with historical crop maps as stratification has proven to be effective and efficient. The estimated soybean area (341,000 km^2 with a SE of 23,000 km^2) was 1.9% higher than the acreage estimate of planted soybeans from the USDA NASS 2016 January estimate and 1.0% lower than the NASS 2015 June survey. A simple random sample without stratification would yield a SE 1.92 times larger than the SE of the implemented stratified sample. A key advantage of the methodology is that the overall area estimate of cultivated soybean could be derived in early September, about two months before harvesting.

The success of the soybean classification was built upon an automated procedure of transforming satellite images to time-series metrics. These metrics can capture the critical features of crop phenology and simultaneously improve the spatiotemporal consistency of the satellite data, both of which are useful for addressing the diversity of soybean cropping at a national scale. Using sample-based area estimates as input to the classification resulted in a calibrated soybean map so that the soybean area of the map matched the sample-based area estimate. The developed method is expected to be tested and applied to other regions in future research.

Acknowledgments

Funding for this research was provided by NASA Land-Cover/Land-Use Change Program (NNX12AC78G and NNX15AK65G) and the Gordon and Betty Moore Foundation. The authors thank Ricardo Aguilar, Chima Okpa, Jacob Noel and Elvis Herrera for assisting with field data collection.

Appendix A. Formulas for estimating accuracy

The estimates summarizing the accuracy of the soybean map were produced using the SURVEYMEANS procedure of the Statistical Analysis Software (SAS version 9.3, SAS Institute Inc., Cary, North Carolina, USA). The sampling design was a two-stage cluster design with the clusters grouped into strata. For the sampling design implemented, the primary sampling unit is a cluster (indicated by the subscript i) and each cluster is assigned to a stratum h. A simple random sample of k_h clusters was selected from the K_h clusters in each stratum h, and within each sampled cluster $n = 10$ pixels were selected via simple random sampling. The general form of the estimator of the three accuracy measures used is a ratio estimator for two-stage cluster sampling within a stratified design (see online documentation provided by SAS version 9.3),

$$\hat{R} = \frac{\sum_{h=1}^{H}\sum_{i=1}^{k_h}\sum_{u=1}^{n} w_{hiu} y_{hiu}}{\sum_{h=1}^{H}\sum_{i=1}^{k_h}\sum_{u=1}^{n} w_{hiu} x_{hiu}} \tag{A1}$$

where u is the index of the sampled pixels ($u = 1, \ldots, n$), i is the cluster index in stratum h ($i = 1, 2, \ldots, k_h$), h is the stratum index ($h = 1, 2, \ldots, H$), x_{hiu} and y_{hiu} are defined to yield the parameter of interest (see below), and $w_{hiu} = K_h/k_h$ is the estimation weight (i.e., inverse of the inclusion probability for the first-stage sample selection) for sample pixel u in cluster i of stratum h. Because all accuracy estimators are ratio estimators, we did not incorporate the second-stage inclusion probability in this analysis because doing so would create the same multiplier in both the numerator and denominator of Eq. (A1) and therefore would not meaningfully impact the estimates (constant multiplier would cancel out in the ratio). The variance estimator for $\hat{R}$ is based on a Taylor series approximation (Särndal et al., 1992):

$$\hat{V}\left(\hat{R}\right) = \sum_{h=1}^{H}\hat{V}_h\left(\hat{R}\right) = \sum_{h=1}^{H}\frac{k_h\left(1-\frac{k_h}{K_h}\right)}{(k_h-1)}\sum_{i=1}^{k_h}\left(g_{hi\cdot}-\bar{g}_{h\cdot\cdot}\right)^2 \tag{A2}$$

where

$$g_{hi\cdot} = \frac{\sum_{u=1}^{n} w_{hiu}\left(y_{hiu}-x_{hiu}\hat{R}\right)}{\sum_{h=1}^{H}\sum_{i=1}^{k_h}\sum_{u=1}^{n} w_{hiu}x_{hiu}} \tag{A3}$$

and

$$\bar{g}_{h\cdot\cdot} = \left(\sum_{i=1}^{k_h} g_{hi\cdot}\right)/k_h. \tag{A4}$$

The variance estimator (Eq. (A2)) does not include the contribution of variance from the subsampling within the first-stage sample clusters.

To estimate overall accuracy, we defined $y_{hiu} = 1$ if pixel u of stratum h in cluster i was correctly classified ($y_{hiu} = 0$ otherwise) and defined $x_{hiu} = 1$ for each pixel sampled. In this case $\hat{R}$ would be the estimated number of correctly classified pixels divided by the estimated total number of pixels, and the estimated ratio would therefore be an estimate of overall accuracy. To estimate user's accuracy of soybean, we defined $y_{hiu} = 1$ if sample pixel u of cluster i from stratum h was correctly classified as soybean (otherwise $y_{hiu} = 0$) and defined $x_{hiu} = 1$ if sample pixel u was classified (mapped) as soybean (otherwise $x_{hiu} = 0$). With these definitions of x and y, $\hat{R}$ estimates the total number of pixels for which both the map and reference class were soybean divided by the estimated total number of pixels mapped as soybean, the definition of user's accuracy of soybean. Lastly, to estimate producer's accuracy of soybean, we defined $y_{hiu} = 1$ if pixel u was correctly classified as soybean (otherwise $y_{hiu} = 0$) and defined $x_{hiu} = 1$ if pixel u had reference class of soybean (otherwise $x_{hiu} = 0$).

Appendix B. Supplementary data

Supplementary data associated with this article can be found in the online version http://dx.doi.org/10.1016/j.catena.2017.01.008. These data include the Google maps of the most important areas described in this article.

References

AAFC, 2015. AAFC Annual Crop Inventory – Data Product Specifications (ISO 19131).

Anderson, M.C., Zolin, C.A., Sentelhas, P.C., Hain, C.R., Semmens, K., Tugrul Yilmaz, M., Gao, F., Otkin, J.A., Tetrault, R., 2016. The evaporative stress index as an indicator of agricultural drought in Brazil: an assessment based on crop yield impacts. Remote Sens. Environ. 174, 82–99.

Becker-Reshef, I., Vermote, E., Lindeman, M., Justice, C., 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. Remote Sens. Environ. 114, 1312–1323.

Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84.

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, 341–358.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Chapman & Hall/CRC, Boca Raton, Florida.

Broich, M., Hansen, M.C., Potapov, P., Adusei, B., Lindquist, E., Stehman, S.V., 2011. Time-series analysis of multi-resolution optical imagery for quantifying forest cover loss in Sumatra and Kalimantan, Indonesia. Int. J. Appl. Earth Obs. Geoinf. 13, 277–291.

Bwangoy, J.-R.B., Hansen, M.C., Roy, D.P., Grandi, G.D., Justice, C.O., 2010. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. Remote Sens. Environ. 114, 73–86.

Carfagna, E., Gallego, F.J., 2005. Using remote sensing for agricultural statistics. Int. Stat. Rev. 73, 389–404.

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.

Chang, J., Hansen, M.C., Pittman, K., Carroll, M., DiMiceli, C., 2007. Corn and soybean mapping in the United States using MODIS time-series data sets. Agron. J. 99, 1654.

Cihlar, J., 2000. Land cover mapping of large areas from satellites: status and research priorities. Int. J. Remote Sens. 21, 1093–1114.

Cochran, W.G., 1977. Sampling Techniques. third ed. John Wiley & Sons, New York.

Cotter, J.J., Tomczack, C.M., 1994. An image analysis system to develop area sampling frames for agricultural surveys. Photogramm. Eng. Remote. Sens. 60, 229–306.

DeFries, R., Hansen, M., Townshend, J., 1995. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. Remote Sens. Environ. 54, 209–222.

Doraiswamy, P.C., Sinclair, T.R., Hollinger, S., Akhmedov, B., Stern, A., Prueger, J., 2005. Application of MODIS derived parameters for regional crop yield assessment. Remote Sens. Environ. 97, 192–202.

FAO, 2015. World program of the census of agriculture 2020, volume I: programme, concepts and definitions. FAO Statistical Development Series. FAO, Rome, p. 204.

Friedl, M.A., Brodley, C.E., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sens. Environ. 61, 399–409.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sens. Environ. 83, 287–302.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., Havlik, P., Mosnier, A., Thornton, P., Wood-Sichra, U., Herrero, M., Becker-Reshef, I., Justice, C., Hansen, M., Gong, P., Abdel Aziz, S., Cipriani, A., Cumani, R., Cecchi, G., Conchedda, G., Ferreira, S., Gomez, A., Haffani, M., Kayitakire, F., Malanding, J., Mueller, R., Newby, T., Nonguierma, A., Olusegun, A., Ortner, S., Rajak, D.R., Rocha, J., Schepaschenko, D., Schepaschenko, M., Terekhov, A., Tiangwa, A., Vancutsem, C., Vintrou, E., Wenbin, W., van der Velde, M., Dunwoody, A., Kraxner, F., Obersteiner, M., 2015. Mapping global cropland and field size. Glob. Chang. Biol. 21, 1980–1992.

Gallego, F.J., 2004. Remote sensing and land cover area estimation. Int. J. Remote Sens. 25, 3019–3047.

Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A., Kussul, O., 2014. Efficiency assessment of using satellite data for crop area estimation in Ukraine. Int. J. Appl. Earth Obs. Geoinf. 29, 22–30.

Gao, B.-c., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 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.

Gonzáles-Alonso, F., Cuevas, J.M., 1993. Remote sensing and agricultural statistics: crop area estimation through regression estimators and confusion matrices. Int. J. Remote Sens. 14, 1215–1219.

Gonzáles-Alonso, F., Soria, S.L., Gozalo, J.M.C., 1991. Comparing two methodologies for crop area estimation in Spain using Landsat TM images and ground-gathered data. Remote Sens. Environ. 35, 29–35.

Gonzalez-Alonso, F., Cuevas, J.M., Arbiol, R., Baulies, X., 1997. Remote sensing and agricultural statistics: crop area estimation in north-eastern Spain through diachronic Landsat TM and ground sample data. Int. J. Remote Sens. 18, 467–470.

Hansen, M., Dubayah, R., Defries, R., 1996. Classification trees: an alternative to traditional land cover classifiers. Remote Sens. Lett. 17, 1075–1081.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R.A., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int. J. Remote Sens. 21, 1331–1364.

Hansen, M.C., Egorov, A., Potapov, P.V., Stehman, S.V., Tyukavina, A., Turubanova, S.A., Roy, D.P., Goetz, S.J., Loveland, T.R., Ju, J., Kommareddy, A., Kovalskyy, V., Forsyth, C., Bents, T., 2014. Monitoring conterminous United States (CONUS) land cover change with Web-Enabled Landsat Data (WELD). Remote Sens. Environ. 140, 466–484.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-centry forest cover change. Science 342, 850–853.

Hansen, M.C., Roy, D.P., Lindquist, E., Adusei, B., Justice, C.O., Altstatt, A., 2008a. A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. Remote Sens. Environ. 112, 2495–2513.

Hansen, M.C., Stehman, S.V., Potapov, P.V., 2010. Quantification of global gross forest cover loss. Proc. Natl. Acad. Sci. U. S. A. 107, 8650–8655.

Hansen, M.C., Stehman, S.V., Potapov, P.V., Loveland, T.R., Townshend, J.R., DeFries, R.S., Pittman, K.W., Arunarwati, B., Stolle, F., Steininger, M.K., Carroll, M., Dimiceli, C., 2008b. Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. Proc. Natl. Acad. Sci. U. S. A. 105, 9439–9444.

Hill, J., Megier, J., 1988. Regional land cover and agricultural area statistics and mapping in the Departement Ardeche, France, by use of Thematic Mapper data. Int. J. Remote Sens. 9, 1573–1595.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.

Johnson, D., Mueller, R., 2010. The 2009 cropland data layer. Photogramm. Eng. Remote Sens. 76, 1201–1205.

Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116–128.

King, L., Adusei, B., Stehman, S., Potapov, P.V., Song, X.-P., Krylov, A., Bella, C.D., Loveland, T.R., Johnson, D.M., & Hansen, M.C. A multi-resolution approach to national-scale cultivated area estimation of soybean. Remote Sens. Environ. (in review)

Lobell, D.B., Asner, G.P., 2004. Cropland distributions from temporal unmixing of MODIS data. Remote Sens. Environ. 93, 412–422.

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. 164, 324–333.

Lohr, S.L., 2010. Sampling: Design and Analysis. second ed. Brooks/Cole, Boston, MA.

Lopresti, M.F., Di Bella, C.M., Degioanni, A.J., 2015. Relationship between MODIS-NDVI data and wheat yield: a case study in northern Buenos Aires province, Argentina. Inf. Process. Agric. 2, 73–84.

Macdonald, R.B., Hall, F.G., 1980. Global crop forecasting. Science 208, 670–679.

Ozdogan, M., 2010. The spatial distribution of crop types from MODIS data: temporal unmixing using independent component analysis. Remote Sens. Environ. 114, 1190–1204.

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, GB1011.

Potapov, P.V., Dempewolf, J., Talero, Y., Hansen, M.C., Stehman, S.V., Vargas, C., Rojas, E.J., Castillo, D., Mendoza, E., Calderón, A., Giudice, R., Malaga, N., Zutta, B.R., 2014. National satellite-based humid tropical forest change assessment in Peru in support of REDD + implementation. Environ. Res. Lett. 9, 124012.

Potapov, P.V., Turubanova, S.A., Hansen, M.C., Adusei, B., Broich, M., Altstatt, A., Mane, L., Justice, C.O., 2012. Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM + data. Remote Sens. Environ. 122, 106–116.

Potapov, P.V., Turubanova, S.A., Tyukavina, A., Krylov, A.M., McCarty, J.L., Radeloff, V.C., Hansen, M.C., 2015. Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. Remote Sens. Environ. 159, 28–43.

Pradhan, S., 2001. Crop area estimation using GIS, remote sensing and area frame sampling. Int. J. Appl. Earth Obs. Geoinf. 3, 86–92.

Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Glob. Biogeochem. Cycles 22, GB1003.

Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., Zhang, C., 2010. Web-enabled Landsat data (WELD): Landsat ETM + composited mosaics of the conterminous United States. Remote Sens. Environ. 114, 35–49.

Särndal, C.E., Swensson, B., Wretman, J., 1992. Model-assisted survey sampling. Springer-Verlag, New York, NY.

Stehman, S.V., 2009. Model-assisted estimation as a unifying framework for estimating the area of land cover and land-cover change from remote sensing. Remote Sens. Environ. 113, 2455–2462.

Stehman, S.V., Wickham, J.D., Smith, J.H., Yang, L., 2003. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: statistical methodology and regional results. Remote Sens. Environ. 86, 500–516.

Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., Cai, X., Vithanage, J., Schull, M.A., Dutta, R., 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, 211–261.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150.

Tyukavina, A., Baccini, A., Hansen, M.C., Potapov, P.V., Stehman, S.V., Houghton, R.A., Krylov, A.M., Turubanova, S., Goetz, S.J., 2015. Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. Environ. Res. Lett. 10, 074002.

Waldner, F., Fritz, S., Di Gregorio, A., Plotnikov, D., Bartalev, S., Kussul, N., Gong, P., Thenkabail, P., Hazeu, G., Klein, I., Löw, F., Miettinen, J., Dadhwal, V., Lamarche, C., Bontemps, S., Defourny, P., 2016. A unified cropland layer at 250 m for global agriculture monitoring. Data 1, 3.

Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: an assessment for the U.S. central Great Plains. Remote Sens. Environ. 112, 1096–1116.

Yan, L., Roy, D.P., 2016. Conterminous United States crop field size quantification from multi-temporal Landsat data. Remote Sens. Environ. 172, 67–86.

You, L., Wood, S., Wood-Sichra, U., Wu, W., 2014. Generating global crop distribution maps: from census to grid. Agric. Syst. 127, 53–60.

Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: a multi-year experiment using Landsat imagery. Remote Sens. Environ. 140, 1–13.