1-2014

Seasonal cultivated and fallow cropland mapping using MODIS-based automated cropland classification algorithm

Zhuoting Wu

Prasad S. Thenkabail

Rick Mueller

Forrest Melton

Lee Johnson

See next page for additional authors

Authors
Zhuoting Wu, Prasad S. Thenkabail, Rick Mueller, Forrest Melton, Lee Johnson, Carolyn Rosevelt, John Dwyer, Jeanine Jones, and James P. Verdin

Journal of
Applied Remote Sensing

Seasonal cultivated and fallow cropland mapping using MODIS-based automated cropland classification algorithm

Zhuoting Wu
Prasad S. Thenkabail
Rick Mueller
Audra Zakzeski
Forrest Melton
Lee Johnson
Carolyn Rosevelt
John Dwyer
Jeanine Jones
James P. Verdin

Seasonal cultivated and fallow cropland mapping using MODIS-based automated cropland classification algorithm

Zhuoting Wu,[a,b] Prasad S. Thenkabail,[a] Rick Mueller,[c] Audra Zakzeski,[c] Forrest Melton,[d,e] Lee Johnson,[d,e] Carolyn Rosevelt,[d] John Dwyer,[f] Jeanine Jones,[g] and James P. Verdin[f]

[a]USGS Western Geographic Science Center, Flagstaff, Arizona 86001
zwu@usgs.gov
[b]Northern Arizona University, Merriam-Powell Center for Environmental Research, Flagstaff, Arizona 86001
[c]USDA NASS, Fairfax, Virginia 22030
[d]CSU Monterey Bay, Seaside, California 93955
[e]NASA Ames Research Center, Moffett Field, California 94035
[f]USGS EROS Data Center, Sioux Falls, South Dakota 57198
[g]California Department of Water Resources, Sacramento, California 94236

Abstract. Increasing drought occurrences and growing populations demand accurate, routine, and consistent cultivated and fallow cropland products to enable water and food security analysis. The overarching goal of this research was to develop and test automated cropland classification algorithm (ACCA) that provide accurate, consistent, and repeatable information on seasonal cultivated as well as seasonal fallow cropland extents and areas based on the Moderate Resolution Imaging Spectroradiometer remote sensing data. Seasonal ACCA development process involves writing series of iterative decision tree codes to separate cultivated and fallow croplands from noncroplands, aiming to accurately mirror reliable reference data sources. A pixel-by-pixel accuracy assessment when compared with the U.S. Department of Agriculture (USDA) cropland data showed, on average, a producer's accuracy of 93% and a user's accuracy of 85% across all months. Further, ACCA-derived cropland maps agreed well with the USDA Farm Service Agency crop acreage-reported data for both cultivated and fallow croplands with *R*-square values over 0.7 and field surveys with an accuracy of ≥95% for cultivated croplands and ≥76% for fallow croplands. Our results demonstrated the ability of ACCA to generate cropland products, such as cultivated and fallow cropland extents and areas, accurately, automatically, and repeatedly throughout the growing season. [DOI: 10.1117/1.JRS.8.083685]

Keywords: automated cropland classification algorithm; MODIS; cultivated croplands; fallow croplands; accuracy assessment; cropland statistics.

Paper 13287 received Jul. 31, 2013; revised manuscript received Oct. 22, 2013; accepted for publication Dec. 31, 2013; published online Jan. 29, 2014.

1 Background

Rapid population growth and ongoing climate change place increasing pressure on food security. The world population is likely to increase to 9.3 billion by 2050 from the current population of 7.073 billion.[1] Such a fast pace of population growth hastens urbanization and industrialization, taking land away from agricultural production, and results in increasing demand for food from a reduced total area of agricultural land. Other factors, such as biofuel production, are shifting demands for food, animal feed, and fuel.[2] Meanwhile, extreme weather events, including severe droughts, are projected to occur more frequently,[3,4] putting even more pressure on water supply and food production. In this context, rapid and timely cropland products (e.g., cropland extent/area, crop type, irrigated or rainfed, and cropping intensities) are of great use to a wide range of end-users including governmental agencies, nongovernmental organizations, farmers, water-use managers, and scientists.

Many previously developed cropland maps have been generated using supervised classification methods that require human intervention or interpretation,[5–8] limiting repeatability of implementation of such methods over large areas and across multiple years. There have already been some earlier attempts at the development of automated or semiautomated methods that can be used to provide maps at regional to national scales on an annual basis using repeatable methods.[9–12] Specifically, an automated cropland classification algorithm (ACCA) approach has been proposed and implemented in Tajikistan.[13,14] The concept of ACCA is to use one or more sources of remotely sensed data as well as any useful secondary data (e.g., precipitation, elevation, and temperature) to produce cropland products. The ACCA is developed upon existing knowledge in order to accurately replicate a reference cropland data layer (CDL) or one or more cropland products. Once the ACCA is developed, it will have the ability to reproduce one or more cropland products routinely and repeatedly for independent time periods (e.g., independent years) using the same type of remote sensing and secondary data used to develop the ACCA.[14]

1.1 *Rationale for This Research: Advances Over Existing State of Knowledge*

Existing cropland mapping methods, including the previously developed ACCA, operate on an annual basis. For example, ACCA reported by Thenkabail and Wu[14] computed irrigated and rainfed croplands extents and areas at the end of the year based on remote sensing data such as Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat and various secondary data. Yet, agricultural food security policies often call for cropland products during various stages of the crop growing season. For instance, during the recent severe drought across the western U.S., there was a constant demand for cropland data throughout the growing season from policy makers, resource managers, and researchers.[15] Increasing climate variability poses further uncertainty in agriculture productivity and seasonality in different regions across the world, and therefore the need for more frequent cropland products throughout the growing season from various end-users cannot be overemphasized. Thereby, the first advance in this research is to develop ACCA to produce seasonal (e.g., month by month during the entire growing season) cropland areas and extents.

Further, even though numerous studies have focused on mapping cultivated cropland (referred to as "cropland" hereafter) across spatial and temporal scales,[6–8,16–21] few have attempted to delineate the extent of fallow croplands (referred to as "fallowland" hereafter). While fallowland statistics are available for some places, it is widely accepted that they have very high uncertainty associated with them, since these fallowland extents/areas have not been systematically studied. The fallowlands consist of complex spectral signatures (e.g., barren, sparse grass/shrub cover, weed are present), have widely varying temporal dynamics, and lack field observations and reference data. Yet, fallowlands have great implications on water use, food production, nutrient cycling, and carbon sequestration and therefore climate change.[22] The agricultural practices on fallowlands have large impacts on soil organic carbon and CO_2 emissions, economic subsidies, and potential environmental impacts, such as erosion, ground water contamination, and trace greenhouse gas productions.[23–25] Mapping of fallowlands and capturing its temporal dynamics throughout the growing season can contribute to an improved understanding of the potential for different management practices to mitigate the negative effects of drought-related cropland fallowing. For example, shortage of water for irrigation and crop production is one of the major impacts of drought in the heavily cultivated Central Valley in California, and therefore timely and accurate information on fallowland acreage is extremely useful in identifying the extent of changes in fallowed acreage due to water shortage during drought and guiding decision making with respect to requests for local water transfers, county drought designations, or state emergency proclamations. Therefore, routine and precise mapping of cropland and fallowland distributions in California has great implications on water use and crop water productivity, which can be achieved in an automated fashion especially under the projected more severe drought for the western U.S.[4] Thereby, the second advance in this research is to build production of fallow cropland extents and areas in the ACCA.

The overarching goal of this study was to develop a seasonal ACCA using MODIS 250-m remote sensing data to produce two unique products: (1) seasonal cropland extents and areas and (2) seasonal fallowland extents and areas. Even though earlier ACCA developed by Thenkabail and Wu[14] computed cropland extents and areas, it was limited to annual products and did not produce fallowlands; both of which are serious limitations. The seasonal ACCA is developed to have automated, rapid, and accurate reproduction capability. State of California was chosen as a pilot region for reasons discussed earlier. Preliminary stakeholder requirements for fallowland acreage monitoring in California indicated that the uncertainty of $\pm 25\%$ may be tolerable (NIDIS California Pilot web conference, 2012), given the current prevailing lack of fallowland monitoring capability. Meanwhile, user requirements for cropland accuracy are vague or are still under development.

Seasonal ACCAs were developed for 5 months throughout the growing season, including June, August, September, October, and December 2012, and tested using independent datasets from June, August, September, and December 2011. Cropland and fallowland mapping accuracies were also evaluated based on field surveys in April, May, June, September, and October 2012. Cropland areas of the 58 counties of California that were derived from the ACCA were also evaluated using independent data sources from farmer-reported surveys [e.g., U.S. Department of Agriculture Farm Service Agency (USDA FSA)]. Given the increasing uncertainty of cropland dynamics in a changing climate, mapping of cropland extent and area on a monthly basis throughout the growing season will be of great value in assessing drought impacts on agricultural production as well as patterns in short- and long-term fallowing of farmland.

Historically, though the use of sensors such as those carried on the Landsat missions can be problematic in areas with high-cloud cover during the growing season, production of fine-resolution cropland maps relied on 1- to 30-m resolution remote sensing imageries. Recent work on cropland classification methods has utilized data from the MODIS instrument, which provides an alternative for monitoring cropland areas [8,17,20,21,26–34] across multiple temporal and spatial scales due to its high-temporal coverage, moderate spatial resolution, and high-quality time-composite products that largely resolve the cloud issues. Thus, in this research, MODIS 250-m data were used to produce seasonal cropland and fallowland extents and areas.

2 Methods

2.1 *Study Area*

The state of California, USA (32°N to 42°N and 114°W to 124°W) was used as the study area to develop, test, and implement a seasonal ACCA. California is one of the most productive agricultural regions in the world, where its agricultural exports play a major role in the state's economy.[35] California produces a variety of crops, some of which are unique commodities grown only in California. According to the USDA National Agricultural Statistics Service (NASS) California Field Office, in 2011, the state's 81,500 farms and ranches generated record \$43.5 billion products revenue, up from the \$38 billion reached during 2010, and therefore California remained the top state in cash farm receipts as it comprised 11.6% of the U.S. total.[36] According to the 2012 NASS CDL for California, about 40,000 km^2 (9.4% of California's total area of 424,000 km^2) was cropland. Farming accounts for about 84% of all human water use in California with the majority of California's cropland being irrigated.[37] The climate is characterized by winter/spring precipitation followed by summer drought. Thus, fallowlands may support weeds or other rainfed volunteer vegetation earlier in the year, but during summer, they are typically very low green vegetation fractions, or barrens, due to the lack of summer rainfall and irrigation.

2.2 *Remote Sensing Data*

An ACCA for California was developed using MODIS 250-m data for the growing season of 2012. Unlike a previous ACCA in Tajikistan, where a combination of MODIS and Landsat data were used,[14] only MODIS time series data were used here, because Landsat Thematic Mapper 5 (TM5) data were discontinued in 2012 and scanline issues associated with the Landsat 7 Enhanced TM (ETM) made the data difficult to use for this application. MODIS normalized difference vegetation index (NDVI) time series have been used widely and proven to be efficient in mapping cropland.[8,17,20,21,27,29–31,34] In this study, MODIS data were solely used to develop the ACCA, and the approach leveraged the frequent temporal coverage provided by MODIS for the seasonal mapping of cropland. In addition, the MODIS-based case study of California in this research can have great implications for implementation of such automated cropland mapping methods to other areas around the world, where higher temporal resolution of MODIS compared with Landsat largely increases the probability of acquiring frequent cloud-free images.

MODIS Terra surface reflectance 8-day composite level 3 (L3) Global 250-m data were obtained through the National Aeronautics and Space Administration (NASA) Earth Observing System Data and Information System (EOSDIS, http://earthdata.nasa.gov/) for the years 2011 and 2012. This dataset is a L3 composite with each pixel containing the best possible L2 observation during an 8-day period, as selected on the basis of high-observation coverage, low-view angle, the absence of clouds or cloud shadow, and aerosol loading.[38] The L3 products have been atmospherically corrected for atmospheric scattering and absorption from atmospheric gases and aerosols, and therefore the resulting surface reflectance is as it would have been measured at ground level. In addition, the L3 products have also been geometrically corrected. All L3 data have been geolocated into a specific map projection, and the gridding process converted the input observation space to the output geometrically correct level.[39] We obtained all the available 8-day composite data from Jan. 1, 2011, to Dec. 31, 2012. Two bands (band 1 and band 2) of the MODIS surface reflectance data were used to calculate NDVI by using the equation: NDVI $= (\text{band } 2 - \text{band } 1)/(\text{band } 2 + \text{band } 1)$, where bands 1 and 2 are red and near-infrared bands, respectively. The NDVI values, which ranged from -1 to $+1$, were then converted to 8-bit scaled NDVI ranging from 0 to 255 by using the equation: Scaled NDVI $= \text{Calculated NDVI} \times 127.5 + 127.5$. Upon generating the scaled NDVI for all MODIS 250-m 8-day composites, a monthly maximum value composite was applied by taking the maximum NDVI value for each pixel from all 8-day composites within the same month, resulting in one maximum MODIS NDVI composite for each month. All monthly maximum MODIS NDVI composites within the same year were then stacked together into a data cube for the ACCA development.[14]

2.3 *Automated Cropland Classification Algorithm Development*

The process of the ACCA development first involves the generation of sets of rules using MODIS NDVI data cube to derive a cropland layer that matches a reference map. For California and the U.S., the USDA annual CDL provides an ideal reference that includes accuracy statistics for all crop classes included in the map. Contrary to the previous ACCA developed in Tajikistan,[14] where the reference cropland layer had to be first created using laborious classification methods, in this study, routinely available CDLs with high accuracies (~80% averaged across crops) were available from NASS for the state of California each year since 2007.

The total cropland areas were comprised of areas cultivated in a given year and areas left fallow (referred to as “fallowland” and looked similar to bare soil or very low-cover vegetation) in the same year. We obtained the year-end CDLs from the NASS CropScape portal (http://nassgeodata.gmu.edu/CropScape/) from 2007 through 2011 to create a 5-year cropland mask including both cultivated and fallow areas, representing the maximum potential cultivated and fallow cropland areas occurred in California within the 5 years, where ACCA can be applied to delineate cropland from noncropland as well as cultivated versus fallow cropland. Within the 5 years, about 20% of the croplands were left fallow in any given year, although the spatial locations of these fallowlands, typically, changed from year to year. We also obtained monthly fallowland data products throughout the growing season from June to December of 2012 from NASS, from which monthly simulated cropland layers (SCLs) were generated as the remaining area from the 5-year cropland mask. Upon resampling the MODIS data cube to 30-m resolution as the SCL, the thresholds of ACCA rules were determined by trial and error, where a threshold of a rule was set when ≤10% error of omission was achieved comparing with the reference SCL. During the ACCA development process, one single rule can only map a certain portion of the total cropland area compared with the reference cropland layer, and additional rules were then written to delineate the unresolved cropland left over from the previous rule sets; therefore, no overlap of cropland area occurred among multiple rule sets. Eventually, all rule sets were compiled together into one single algorithm when over 90% of the total cropland from the ACCA-derived cropland layer matched with the reference cropland layer pixel-by-pixel. The seasonal ACCA delineated actively cultivated areas and left aside the fallowlands within the total cropland areas. Thus, both cropland and fallowlands were established throughout the growing season. All algorithms coding and testing were conducted in the ERDAS Imagine 2011 Modeler.

2.4 *Accuracy Assessment*

In this research, three unique approaches of accuracy assessments were adopted to ensure the strength of the seasonal ACCA.

2.4.1 *Simulated cropland layers from USDA NASS: pixel-based error matrices between remote sensing products and reference maps*

The SCLs derived from the USDA NASS cropland products for the growing season of 2012 were available for months of June, August, September, October, and December, which were used to generate pixel-by-pixel error matrices[40] for comparison with ACCA-derived cropland layers. The advantage of ACCA is to map cultivated and fallowed agricultural lands automatically once they are developed, so we obtained the retrospective SCLs of California for June, August, September, and December 2011 as independent data layers to test the applicability of the ACCA developed using SCLs of 2012.

2.4.2 *Farm Service Agency crop acreage data: comparisons between areas derived from remote sensing products and FSA census data*

Monthly county-level crop acreage data from the USDA FSA were available starting from August 2011. The crop acreage data were based on individual farmers reporting and were aggregated at the county level. Contrary to the pixel-by-pixel based error matrices above (Sec. 2.4.1), this dataset permits an area-based accuracy assessment by comparing the area statistics between the ACCA-derived cropland layers and the crop acreage data. We obtained the monthly crop acreage data throughout the growing season starting from August 2011 to December 2012 to assess the ability of ACCA to generate cropland area statistics that agree with the county-level reporting data from FSA.

2.4.3 *Ground-based field surveys: field data validating remote sensing products*

In the early growing season of 2012, 10 east-west field transects were established in California, spanning horizontally across the major cropland areas in California (Fig. 1). The transects ranged in distance from 20 to 54 km across the Central Valley. Data were collected for every field along each transect based on visual inspection of each field at a location on the edge of the field adjacent to an access road. Digital photos were taken for each field, and descriptive information was recorded including geographic location (latitude and longitude), initial subjective classification (e.g., crop and idle at present time), bare soil condition (e.g., present or not, tilled, beds shaped, irrigated wet soil, and flooded), weed condition (e.g., present or not, color, fractional cover, height, and type), cover crop condition (e.g., present or not, color, fractional cover, height, and type), and crop condition (e.g., present or not, type, structure, color, and condition, including emergent, growing, senescent, residue, recently harvested, fractional cover, height, and irrigation system type). Data were entered into a geographic information system database that linked georegistered field locations and associated attributes collected during the field surveys. Two early season field surveys were conducted for each transect (April/May and May/June) and sampled

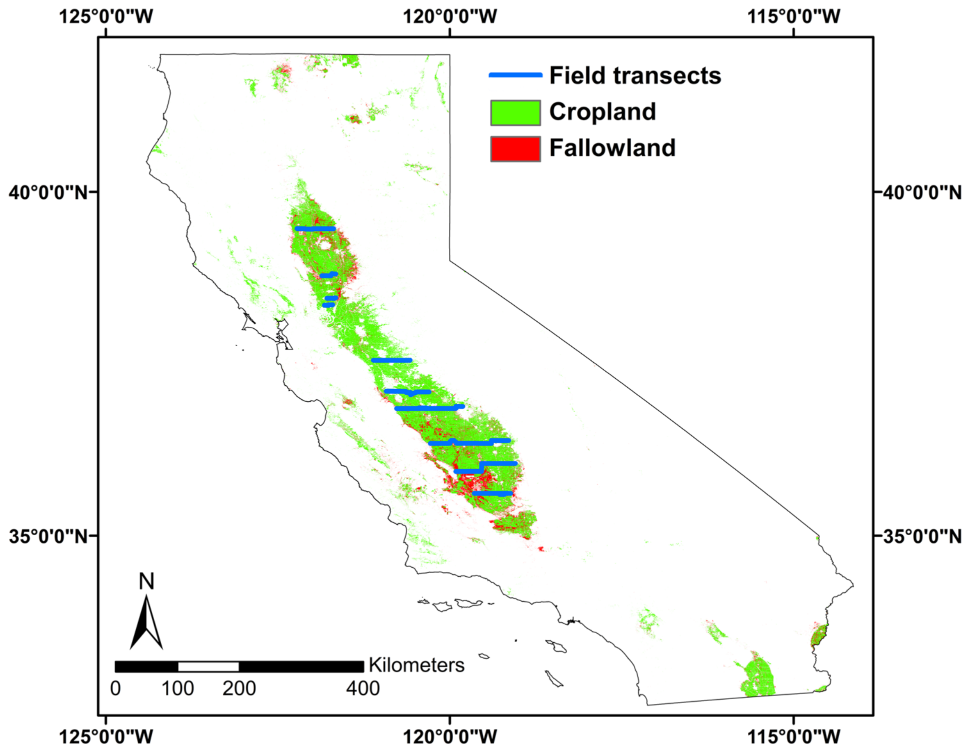

Fig. 1 Early seasons field transects for mapping cropland and fallowland in California. The cropland and fallowland maps are from the NASS Cropland Data Layer (CDL) of 2012. Late season sampling was concentrated in the fallowland region near 35N/120W.

all fields along each transect route. Since the original surveys contained a very small percentage of fields that did not have a clearly established crop, two late season field surveys were conducted in September and October. Due to logistical constraints, these surveys focused on a subset of fields identified as fallow by either the NASS or ACCA fallowland maps for August 2012. A general principle of ground data collection for accuracy assessment was to collect 100 samples per classification category (e.g., cultivated or fallow cropland) when dealing with a large area.[41] In order to meet such sampling criteria for accuracy assessment, we sampled over 1500 fields across the 10 transects throughout the growing season of 2012. We used the early season field surveys to assess the accuracy of ACCA-derived cropland map, and late season field surveys to validate ACCA-derived fallowland map of 2012.

3 Results

3.1 *Seasonal ACCA for Cropland of California*

Seasonal ACCAs were developed throughout the growing season of 2012 for the months of June, August, September, October, and December based on the reference SCLs for the corresponding months (e.g., Fig. 2 illustrates the seasonal ACCA for the month of August 2012). The algorithm was designed to detect the presence of crops, as indicated by elevated NDVI or standalone near-infrared reflectance. The algorithm development process started with using the yearly total NDVI to delineate cropland, followed by using NDVI from critical months out of the year to map seasonal croplands (e.g., Fig. 2). The algorithm for each month followed a similar structure, but the thresholds of the rule sets varied as the growing season progressed and more monthly MODIS NDVI data layers were incorporated into the dataset. The ACCA for the month of August (Fig. 2), for example, involved MODIS NDVI data for the months of January through July. As the growing season progressed, more monthly MODIS NDVI data layers were used in the algorithm in the subsequent months to more rigorously map cropland.

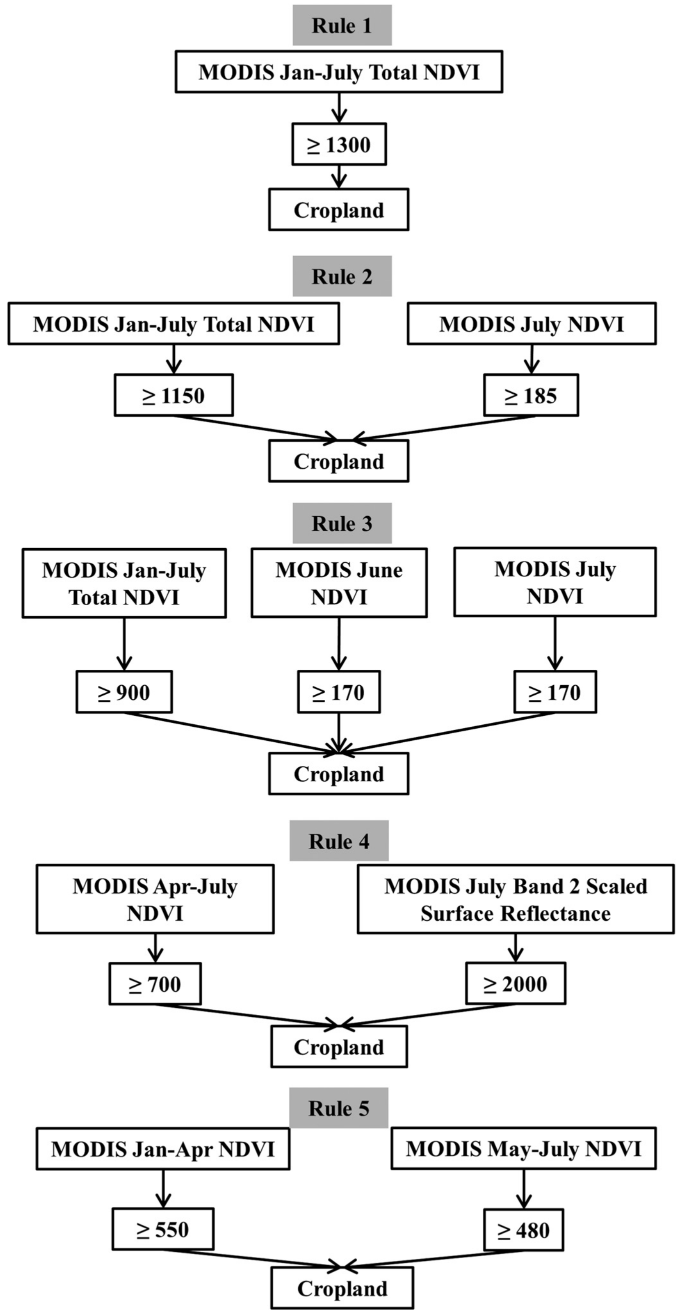

Fig. 2 Sample automated cropland classification algorithm (ACCA) of August 2012 for the state of California. Total NDVI is the additive sum of monthly NDVI, and MODIS surface reflectance has a scale factor of 0.0001.

In addition, the band 2 (near-infrared) surface reflectance during July was also used in the algorithm (except for June ACCA). During the algorithm development process, the cumulative monthly NDVI was the most effective in delineating the largest amount of cultivated area and set as Rule 1 (Fig. 2). In the August ACCA (Fig. 2), for example, Rule 1 using monthly total NDVI from January to July delineated 56% of the total cropland area. Rules 2 and 3 used monthly total NDVI as well as individual monthly NDVI of critical months during the growing season to delineate additional cropland area that was not captured by Rule 1 (Fig. 2). Rules 4 and 5 used seasonal (e.g., April to July and January to April) cumulative NDVI as well as MODIS band 2 surface reflectance of July to delineate the remaining cropland (Fig. 2).

3.2 *ACCA-Derived Cropland Maps*

We applied the seasonal ACCAs (June, August, September, October, and December) to the MODIS monthly NDVI data cube of 2012 (Sec. 2.2) and generated cropland layers for the corresponding months. To avoid repetition, only cropland output for August was illustrated with its algorithm shown in Fig. 2. The ACCA-derived cropland map for August 2012 agreed well with the August SCL, demonstrated by an overwhelmingly overlapped cropland area captured both in the ACCA-derived cropland map and SCL [Fig. 3(a)]. There were some fractional areas that were captured only by ACCA in the northern Central Valley, and some other isolated areas only present in SCL in the southern Central Valley [Fig. 3(a)]. The August ACCA (Fig. 2) was then applied on an independent MODIS monthly NDVI data cube from 2011, and the output was compared with the August SCL of 2011 [Fig. 3(b)]. Similar to the one in 2012 [Fig. 3(a)], the ACCA-derived cropland map of August 2011 showed substantial agreement with SCL of August 2011 [Fig. 3(b)]. Compared with the results from August 2012 [Fig. 3(a)], there were slightly more cropland areas that were only present in the ACCA-derived cropland map in the northern Central Valley and very minimal cropland areas that were solely in the SCL throughout the state in the ACCA-derived cropland map of August 2011[Fig. 3(b)].

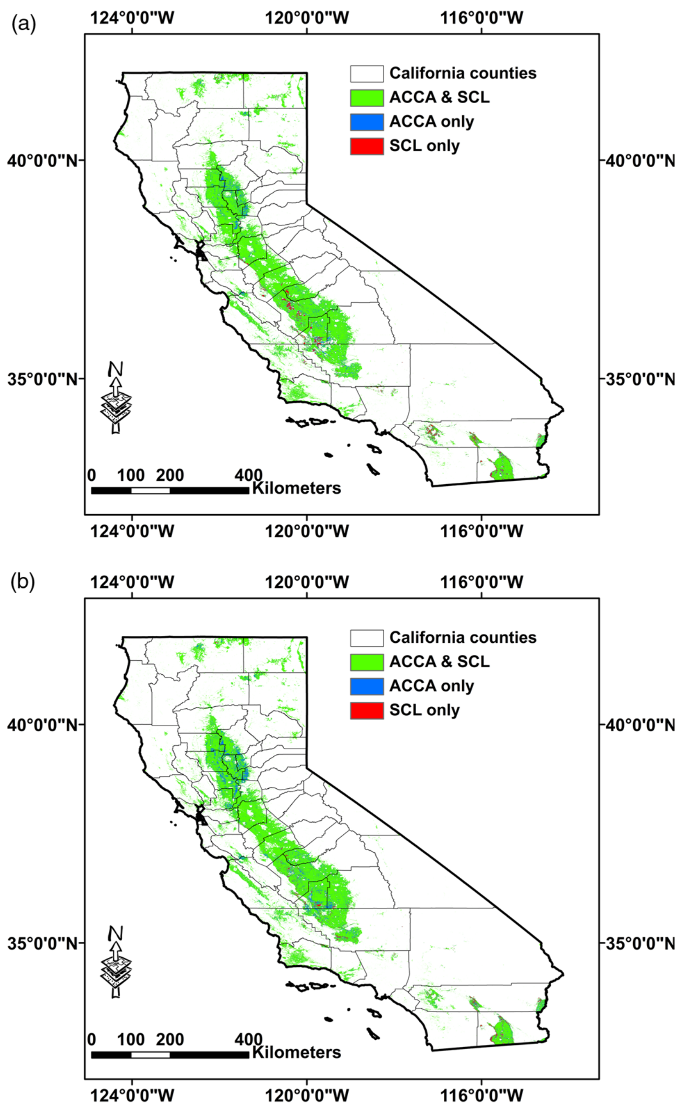

Fig. 3 Comparison of ACCA-derived cropland maps and SCLs of California for (a) August 2012 and (b) August 2011. Common cropland areas in both ACCA-derived croplands and SCLs, areas only in ACCA-derived croplands, and areas only in SCLs were shown here.

Table 1 Error matrix (number of pixels) from the comparison between ACCA-derived cropland layer of August 2012 versus SCL of August 2012.

		ACCA-derived				
		Cropland	Noncropland	Row total	Producer's accuracy	Errors of omission
S C L	Cropland	37185270	2368473	39553743	94%	6%
	Noncropland	7186236	366236781	373423017	98%	2%
	Column total	44371506	368605254			
	User's accuracy	84%	99%			
	Errors of commission	16%	1%			
				Overall accuracy	98%	
				K_{hat}	0.9	

3.3 *Accuracy Assessment for ACCA-Derived Cropland Maps*

3.3.1 *Pixel-based error matrix*

Pixel-by-pixel based comparison between the ACCA-derived cropland map for August 2012 and the SCL for August 2012 showed a producer's accuracy of 94% and a user's accuracy of 84% for cropland (Table 1). The ACCA performed very well in delineating cropland from noncropland, yet there is still mixture of other land-use/land-cover classes in the ACCA-derived cropland map.

Throughout the growing season in 2012, producer's accuracies were above 90% and user's accuracies were above 84% for croplands across all months (Table 2). The ACCA of December used all 12 monthly MODIS NDVI data layers in 2012, and therefore it performed best in resolving the confusion of other land-cover /land-use classes with cropland, with the highest user's accuracy of 93% (Table 2). Overall, the performance of ACCA was consistent on a monthly basis throughout the growing season of 2012, for which the algorithm was developed.

The strength of building an ACCA is to apply it to independent data layers and automatically generate cropland products akin to the one for which the algorithm was developed. The seasonal ACCAs developed for the months of 2012 were applied on the corresponding months of an independent year including June, August, September, and December 2011. Throughout the growing season, the producer's accuracies were above 92%, which were slightly higher than those in 2012 (Table 3). The user's accuracies were above 82% across all months during the growing season of 2011 (Table 3). Thus, ACCA can produce cropland maps on a monthly basis throughout the growing season even for an independent year with high level of accuracies, exhibiting the capability of the seasonal ACCAs to accurately compute cropland extent for various months throughout the growing season year after year.

Table 2 Summary of error matrices of ACCA-derived seasonal cropland layers versus SCLs of California for the growing season of 2012.

Month	Producer's accuracy (%)	User's accuracy (%)	Errors of omission (%)	Errors of commission (%)
June	92	86	8	14
August	94	84	6	16
September	92	85	8	15
October	92	85	8	15
December	90	93	10	7

Table 3 Summary of error matrices of ACCA-derived seasonal cropland layers versus SCLs of California for the growing season of 2011.

Month	Producer's accuracy (%)	User's accuracy (%)	Errors of omission (%)	Errors of commission (%)
June	93	83	7	17
August	96	82	4	18
September	92	83	8	17
December	92	82	8	18

3.3.2 *Area-based comparison with ground truth data at the county level*

USDA FSA has produced county-level crop acreage data on a monthly basis throughout the growing season since August 2011. County-level cropland area was compared between the ACCA-derived cropland map and the FSA county crop acreage data. We illustrated here the linear regression between ACCA-derived cropland area and crop acreage data for December 2012 for the 58 counties of California with a slope of 1 [Fig. 4(a)]. Linear regressions showed that the ACCA agreed well with the crop acreage data over the growing season (Table 4), although ACCA tended to overestimate the cultivated area for September and October 2012 (Table 4), since not all farmers participate in the crop acreage data collection program throughout the growing season. We also compared ACCA-derived fallowland areas to the farmers reported fallow acreage at the county level in 2012. The vast majority of the fallowland of California was located within the Central Valley, where fallowland acreage was consistently reported in the FSA

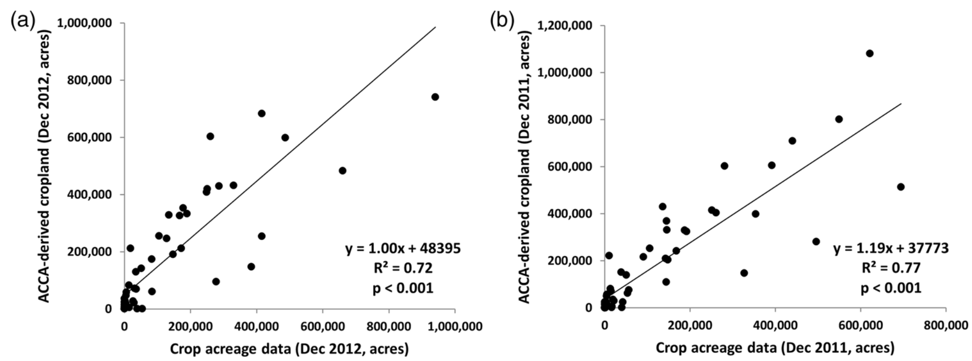

Fig. 4 The ACCA-derived county-level cropland areas of (a) December 2012 and (b) December 2011 for California, in comparison with cropland areas derived from crop acreage data from the Farm Service Agency (FSA) for the corresponding months.

Table 4 Summary of linear regressions of ACCA-derived county-level cropland and fallowland (Central Valley counties) areas against the crop acreage data from the FSA throughout the growing season of 2012 (year during which ACCA was developed) and 2011 (independent year).

	Cropland 2012		Cropland 2011		Fallowland 2012		Fallowland 2011	
Month	Slope	*R*-square	Slope	*R*-square	Slope	*R*-square	Slope	*R*-square
August	1.07	0.70	1.35	0.75	1.24	0.72	3.37	0.85
September	1.26	0.70	1.24	0.76	2.38	0.77	3.15	0.77
October	1.46	0.72	1.22	0.77	2.54	0.79	3.07	0.79
December	1.00	0.72	1.19	0.77	2.62	0.75	3.39	0.80

dataset. R-square values of linear regressions between ACCA-derived fallowland areas and FSA crop acreage data fallowland areas during the growing season of 2012 were all above 0.72 and the slopes were above 1, indicating that the ACCA overestimated the fallowland area consistently (Table 4).

To test the repeatability of ACCA-generated cropland area statistics, we applied the ACCA of December 2012 to the MODIS data cube of 2011 to derive a cropland map of December 2011 and compared the county-level statistics of ACCA-derived cropland area of December 2011 versus the crop acreage data of December 2011 from the FSA. The linear regression showed an R-square of 0.77, and ACCA-derived cropland area tended to overestimate the reported county-level cropland area from the farmers [Fig. 4(b)]. Throughout the growing season of 2011, the county-level cropland area statistics agreed well with the crop acreage data with the R-square values of the linear regressions above 0.75 for all months (Table 4). The county-level fallowland areas showed a good agreement with the FSA crop acreage data fallow area with R-square values of linear regressions over 0.77 across all months. Similar to the results from 2012, ACCA tended to overestimate the cultivated and fallow area with the slopes above 1 (Table 4), which could be due to the data collection process of the crop acreage data, where not all farmers were included. For cropland, as the growing season progressed from August to December in 2011, ACCA more accurately captured the cropland area from the crop acreage data, demonstrated by a progressively smaller slope (closer to 1) and larger R-square value (Table 4).

3.3.3 *Field survey-based accuracy assessment*

Two groups of field surveys were conducted throughout the growing season of 2012 including the early season (April to June) and late season (September to October). We used the early season "crop-present" survey fields to assess the accuracy of the ACCA-derived cropland map of August 2012, and late season "fallow" fields to assess the accuracy of the ACCA-derived fallowland map of October 2012. The early season field surveys covered the heavily cultivated Central Valley area including all major crops [Fig. 5(a)]. The ACCA-derived cropland map for August 2012 showed a very high producer's accuracy of 95% and user's accuracy of 96% (Table 5). The ACCA-derived fallowland map for October 2012 agreed well with the field survey with a producer's accuracy of 84% and user's accuracy of 76% (Table 5) based on 210 surveyed fields, where higher concentrations of fallowland was located in California [Fig. 5(b)]. Greater uncertainty in fallowland extent compared with the cultivated area was expected, given its complex spectral composition and temporal dynamics. Some level of class confusion between cultivated and fallow croplands resulted from spectral insensitivity. For example, newly established plantations on a previously fallowland [see the bottom right field survey photo with a zoom-in inset in Fig. 5(b)] typically have very low fractional cover and were misclassified as fallowland by the ACCA, due to the low proportion of green vegetation. Conversely, some fallowland fields identified in the field surveys had sparse to medium levels of green shrub cover [see the upper right field survey photo in Fig. 5(b)], whereas ACCA defined fallowland as cropland areas exhibiting bare soil-like spectral signatures (lower NDVI). Furthermore, agriculture in California is characterized by homogenous cultivated fields [Fig. 5(a)], and therefore the MODIS-based ACCA worked very well in mapping cultivated area, but the relatively coarse spatial resolution of MODIS data compared with the survey fields may not be sufficient to delineate smaller and/or isolated fallowlands. Yet, ACCA can still provide valuable information for assessing drought impacts in agricultural regions with accuracy standard of $\pm 25\%$ or better for mapping fallowland extent for various months within the growing season.

4 Discussion

The process of developing a seasonal ACCA for a region was demonstrated and implemented to derive cropland maps and crop area statistics from independent datasets automatically. The ACCA-derived products for cropland compare favorably with SCLs and ground surveys and meet established accuracy standards for croplands and fallowlands. The run time for the ACCA for California was approximately 1 h on a Dell Precision T7500 desktop to produce

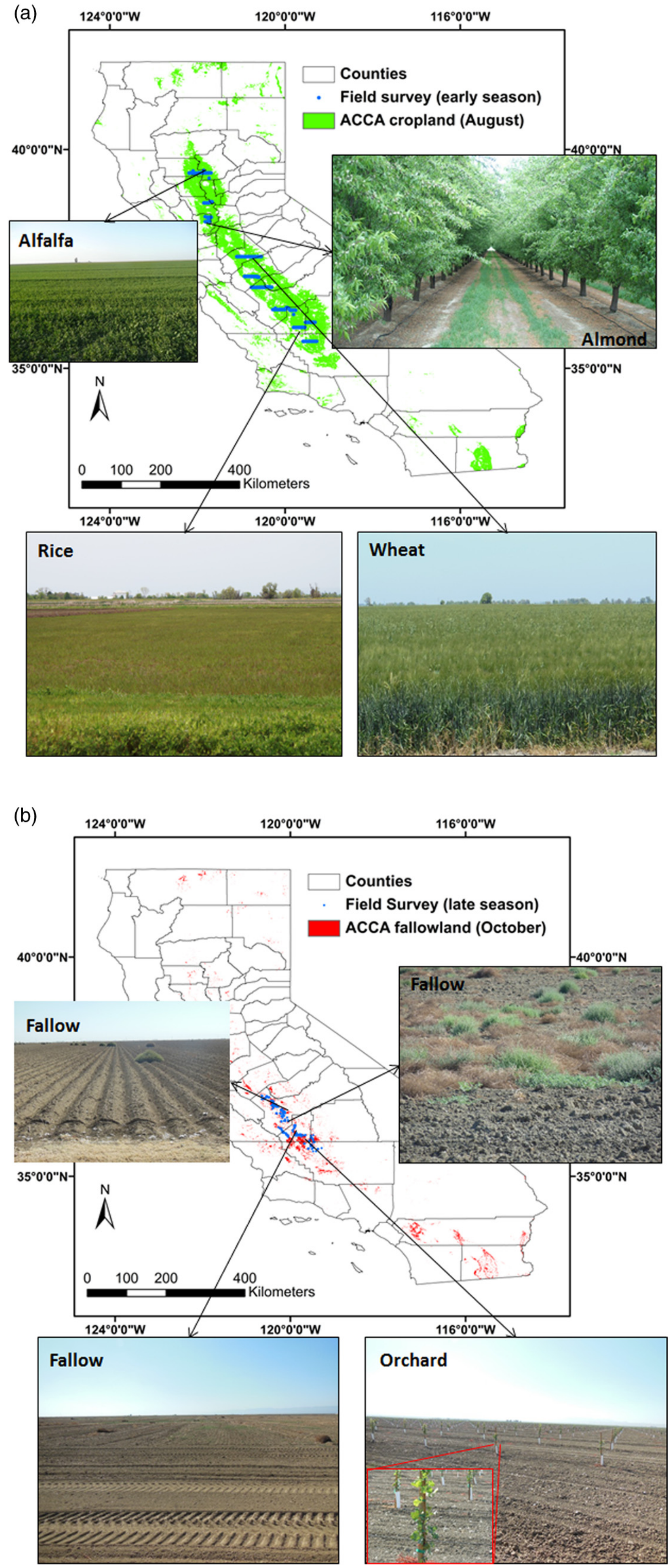

Fig. 5 Comparison of (a) early season (April to June) and (b) late season (September to October) field surveys with ACCA-derived (a) cropland map of August and (b) fallowland map of October 2012. Example field survey photos for cultivated and fallow fields are presented.

Table 5 Summary of field survey-based accuracy assessments (# of fields) of ACCA-derived cropland and fallowland maps of 2012.

		ACCA-derived cropland (August 2012)		
		Crop	Total	Producer's accuracy
Field survey (early season)	Crop-present	1233	1305	95%
	Total	1278		
	User's accuracy	96%		
		ACCA-derived fallowland (October 2012)		
		Fallow	Total	Producer's accuracy
Field survey (late season)	Fallow	160	191	84%
	Total	210		
	User's accuracy	76%		

cropland and fallowland maps and area statistics, once the remote sensing datasets were composed. The seasonal ACCA approach can be fully automated, allowing it to be applied operationally to MODIS data to generate maps of croplands and fallowlands for the corresponding months of independent years with good agreement to reference datasets, once it was developed and tested. This makes the ACCA approach very useful for generating time series either on a yearly basis or monthly basis throughout the growing season, as we demonstrated in this research. For example, NASS CDLs were not available before 2007 nor were monthly county-level statistics available from FSA before 2011; therefore, ACCA can be used to generate cropland maps and county-level cropland areas to assess past drought impacts as well as to produce monthly data products during drought events with a minimal time-lag to inform decision making. More back-testing and calibration through the past years can improve the model and adjust model parameters to account for climatic variability across years.

Another advantage of ACCA is the flexibility to produce either spatially explicit cropland maps or nonspatially explicit cropland area statistics (e.g., at the county level) throughout the growing season. Although other routine cropland maps, such as the NASS annual CDLs, were available, such data are only available at the end of the calendar year and are not available during the growing season. NASS does not release intermediate crop-specific CDL products for June, August, September, or October CDL, as they were considered preliminary, market sensitive, and confidential. Thus, the best and final products were released to the public on an annual basis, which were produced at the end of the growing season, when the majority of the ground data and satellite imagery was available for processing, and thus provided the most accurate final land-cover classification product. On the contrary, a seasonal ACCA on a monthly basis throughout the growing season can provide rapid and accurate cropland maps and statistics; especially such processes can be automatic and labor efficient. Furthermore, the concept of ACCA can be applied to other regions of the world, where cropland data are unreliable or unavailable altogether.[7,42]

In addition, fallowland products can be created after generating ACCA-derived cropland maps. Fallowland areas were either not reported in most cropland products or reported with high uncertainty,[8,21,27,31,43–45] mostly due to its temporal dynamics and mixed signatures from multiple types of fallowlands. Yet fallow acreage needs to be timely updated throughout the growing season to support state or regional decision making for local water transfer requests, or even emergency assistance in extreme drought cases, especially given the projected more frequent, prolonged, and severe droughts under current climate change.[3] One of the strengths of ACCA is that it did not attempt to directly classify fallowland, hence avoid these complications, and was able to achieve relatively high accuracy and meet the accuracy requirement for fallowland.

From the pixel-, area-, farmers-, and field survey-based accuracy assessments, ACCA-derived cropland and fallowland products were proven to be accurate and comparable to other remote sensing-derived (i.e., NASS SCLs) and ground-based (i.e., FSA crop acreage data) cropland products, indicating that the ACCA-derived cropland maps well captured the temporal dynamics of cropland area during the growing season and are likely useful for crop growth and water consumption monitoring and food security decision making. Yet, challenges were apparent to accurately map fallowland during the growing season. For example, confusions between fallowland and newly established plantation in the cultivated land could be reduced by using modified vegetation indices that minimizes the influence of soil background reflectance such as a Soil-Adjusted Vegetation Index,[46] Modified Soil-Adjusted Vegetation Index,[47] and Optimized Soil-Adjusted Vegetation Index,[48] and computation of such indices from the newly launched Landsat 8 at a finer resolution can improve the mapping of isolated and fragmented fallowland. Thus, clear definition and additional knowledge on the characteristics of the fallowland as well as the usage of high-resolution remote sensing imagery and modified vegetation indices can improve and strengthen the ACCA to produce accurate seasonal fallowland maps.

The accuracy of ACCA is often tightly tied to the accuracy of the reference CDLs, and therefore a strong knowledge base is essential for developing ACCA for any particular geographic region and expanding ACCA to broader spatial scales. Thus, the concept of ACCA can be applied to other geographic regions, where reference cropland maps are available or can be produced. Algorithm developed in California cannot be directly applied to other regions, and rule sets need to be modified following the similar structure of existing ACCA and tested before an ACCA can be implemented. Furthermore, the data cube compiled to run ACCA can also differ from region to region. For example, with the launch of Landsat 8 in 2013, high-quality Landsat data become available and can be very useful in delineating cropland in more heterogeneous areas, taking advantage of its high-spatial resolution (30 m) and relatively high-temporal resolution (16 days). Meanwhile, additional remote sensing data and/or secondary data may be required in topographically complex regions.[14]

5 Conclusion

This research demonstrated the development of a seasonal ACCA using MODIS 250-m data for the state of California to generate cropland and fallowland extents and areas for various months throughout the growing season automatically, rapidly, and accurately. The ability of seasonal ACCA to produce accurate results for independent datasets was well established with three approaches. First, when compared with USDA NASS SCL reference datasets, the ACCA showed producer's accuracies ≥90% and user's accuracies ≥82% for croplands. Second, when compared with FSA crop acreage data, the ACCA explained ≥70% variability in cultivated areas of the 58 California counties with the slopes converging toward 1 with the progression of the season. Third, the ACCA also explained ≥72% variability in fallow areas of the Central Valley of California. The ACCA demonstrated ≥95% accuracies for cropland and ≥76% accuracies for fallowland compared with field surveys. The results clearly imply the ability of seasonal ACCA to automatically and accurately generate cropland and fallowland extents and areas for the past and future years, once the remote sensing data cube, akin to the one used for model development, of the months/years of interest is compiled. The ACCA concept can be potentially applied to any other geographic regions in the world, where the ability of ACCA to automatically, accurately, and rapidly generate cropland extents and areas repeatedly year after year and during various months within a year will help cropland water-use assessments, water productivity studies, food security analyses, and decision making during droughts.[49]

Acknowledgments

This work is supported by the U.S. Geological Survey's (USGS) WaterSMART (Sustain and Manage America's Resources for Tomorrow) project and Famine Early Warning Network (FEWSNET) project. Inputs on algorithm development from the team members of the

USGS Powell Center working group on global croplands (http://powellcenter.usgs.gov/current_projects.php#GlobalCroplandMembers) are deeply appreciated. Comments from the USGS internal reviewers Kristin Byrd and Laura Norman are greatly appreciated. Funding support from USGS Land Change Science (LCS) and Land Remote Sensing (LRS) programs are gratefully acknowledged. Support of the Western Geographic Science Center is deeply appreciated. The use of trade, product, or firm names is for descriptive purposes only and does not constitute endorsement by the U.S. government.

References

1. United Nations Department of Economic and Social Affairs Population Division, World Population Prospects, the 2010 Revision, http://esa.un.org/wpp (2011).
2. T. Searchinger et al., "Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change," *Science* **319**(5867), 1238–1240 (2008), http://dx.doi.org/10.1126/science.1151861.
3. Intergovernmental Panel on Climate Change (IPCC), *Climate change 2007: synthesis report, contribution of working groups I to the fourth assessment report of the intergovernmental panel on climate change*, Geneva, Switzerland, R. K. Pachauri and A. Reisigner, Eds., pp. 104 (2007).
4. R. Seager and G. A. Vecchi, "Greenhouse warming and the 21st century hydroclimate of southwestern North America," *Proc. Natl. Acad. Sci. U. S. A.* **107**(50), 21277–21282 (2010), http://dx.doi.org/10.1073/pnas.0910856107.
5. I. A. EL-Magd and T. W. Tanton, "Improvements in land use mapping for irrigated agriculture from satellite sensor data using a multi-stage maximum likelihood classification," *Int. J. Remote Sens.* **24**(21), 4197–4206 (2003), http://dx.doi.org/10.1080/0143116031000139791.
6. C. M. Biradar et al., "A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing," *Int. J. Appl. Earth Obs. Geoinf.* **11**(2), 114–129 (2009), http://dx.doi.org/10.1016/j.jag.2008.11.002.
7. P. S. Thenkabail et al., "Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium," *Int. J. Remote Sens.* **30**(14), 3679–3733 (2009), http://dx.doi.org/10.1080/01431160802698919.
8. K. Pittman et al., "Estimating global cropland extent with multi-year MODIS data," *Remote Sens.* **2**(7), 1844–1863 (2010), http://dx.doi.org/10.3390/rs2071844.
9. S. K. Maxwell et al., "An automated approach to mapping corn from Landsat imagery," *Comput. Electron. Agric.* **43**(1), 43–54 (2004), http://dx.doi.org/10.1016/j.compag.2003.09.001.
10. A. Kumar, S. K. Ghosh, and V. K. Dadhwal, "ALCM: automatic land cover mapping," *J. Indian Soc. Remote Sens.* **38**(2), 239–245 (2010), http://dx.doi.org/10.1007/s12524-010-0030-x.
11. J. Xia, J. Liu, and G. Wang, "A novel automatic classification method based on multi-spectrum remote sensing imagery," in *Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC)*, Harbin, China, pp. 1383–1386 (2011).
12. Y. Pan et al., "Mapping cropland distributions using a hard and soft classification model," *IEEE Trans. Geosci. Remote Sens.* **50**(11), 4301–4312 (2012), http://dx.doi.org/10.1109/TGRS.2012.2193403.
13. P. S. Thenkabail et al., "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 (2010), http://dx.doi.org/10.3390/rs2010211.
14. P. S. Thenkabail and Z. Wu, "An automated cropland classification algorithm (ACCA) for Tajikistan by combining Landsat, MODIS, and secondary data," *Remote Sens.* **4**(10), 2890–2918 (2012), http://dx.doi.org/10.3390/rs4102890.
15. F. Kogan, T. Adamenko, and W. Guo, "Global and regional drought dynamics in the climate warming era," *Remote Sens. Lett.* **4**(4), 364–372 (2013), http://dx.doi.org/10.1080/2150704X.2012.736033.

16. J. Chang et al., "Corn and soybean mapping in the United States using MODN time-series data sets," *Agron. J.* **99**(6), 1654–1664 (2007), http://dx.doi.org/10.2134/agronj2007.0170.
17. B. D. Wardlow, S. L. Egbert, and J. H. Kastens, "Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains," *Remote Sens. Environ.* **108**(3), 290–310 (2007), http://dx.doi.org/10.1016/j.rse.2006.11.021.
18. G. J. Husak et al., "Crop area estimation using high and medium resolution satellite imagery in areas with complex topography," *J. Geophys. Res. Atmos.* **113**(D14), D14112 (2008), http://dx.doi.org/10.1029/2007JD009175.
19. P. S. Thenkabail et al., *Remote Sensing of Global Croplands for Food Security*, CRC Press—Taylor and Francis Group, Boca Raton (2009).
20. V. Dheeravath et al., "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 (2010), http://dx.doi.org/10.1016/j.isprsjprs.2009.08.004.
21. Y. Shao et al., "Mapping cropland and major crop types across the Great Lakes Basin using MODIS-NDVI data," *Photogramm. Eng. Remote Sens.* **76**(1), 73–84 (2010), http://dx.doi.org/10.14358/PERS.76.1.73.
22. R. Lal, "Carbon sequestration in dryland ecosystems," *Environ. Manage.* **33**(4), 528–544 (2004), http://dx.doi.org/10.1007/s00267-003-9110-9.
23. A. Bauer and A. L. Black, "Soil carbon, nitrogen, and bulk density comparisons in two cropland tillage systems after 25 years and in virgin grassland," *Soil Sci. Soc. Am. J.* **45**(6), 1166–1170 (1981), http://dx.doi.org/10.2136/sssaj1981.03615995004500060032x.
24. J. D. Reeder, G. E. Schuman, and R. A. Bowman, "Soil C and N changes on conservation reserve program lands in the Central Great Plains," *Soil Tillage Res.* **47**(3–4), 339–349 (1998), http://dx.doi.org/10.1016/S0167-1987(98)00122-6.
25. J. S. Janosky, D. L. Young, and W. F. Schillinger, "Economics of conservation tillage in a wheat–fallow rotation," *Agron. J.* **94**(3), 527–531 (2002), http://dx.doi.org/10.2134/agronj2002.0527.
26. P. C. Doraiswamy et al., "Crop condition and yield simulations using Landsat and MODIS," *Remote Sens. Environ.* **92**(4), 548–559 (2004), http://dx.doi.org/10.1016/j.rse.2004.05.017.
27. D. B. Lobell and G. P. Asner, "Cropland distributions from temporal unmixing of MODIS data," *Remote Sens. Environ.* **93**(3), 412–422 (2004), http://dx.doi.org/10.1016/j.rse.2004.08.002.
28. P. S. Thenkabail, M. Schull, and H. Turral, "Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data," *Remote Sens. Environ.* **95**(3), 317–341 (2005), http://dx.doi.org/10.1016/j.rse.2004.12.018.
29. T. W. Biggs et al., "Irrigated area mapping in heterogeneous landscapes with MODIS time series, ground truth and census data, Krishna Basin, India," *Int. J. Remote Sens.* **27**(19), 4245–4266 (2006), http://dx.doi.org/10.1080/01431160600851801.
30. M. Ozdogan and G. Gutman, "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 (2008), http://dx.doi.org/10.1016/j.rse.2008.04.010.
31. B. D. Wardlow and S. L. Egbert, "Large-area crop mapping using time-series MODIS 250 m NDVI data: an assessment for the US Central Great Plains," *Remote Sens. Environ.* **112**(3), 1096–1116 (2008), http://dx.doi.org/10.1016/j.rse.2007.07.019.
32. R. Nemani et al., "Monitoring and forecasting ecosystem dynamics using the Terrestrial Observation and Prediction System (TOPS)," *Remote Sens. Environ.* **113**(7), 1497–1509 (2009), http://dx.doi.org/10.1016/j.rse.2008.06.017.
33. R. S. Lunetta et al., "Monitoring agricultural cropping patterns across the Laurentian Great Lakes Basin using MODIS-NDVI data," *Int. J. Appl. Earth Obs. Geoinf.* **12**(2), 81–88 (2010), http://dx.doi.org/10.1016/j.jag.2009.11.005.
34. M. K. Gumma et al., "Mapping rice areas of South Asia using MODIS multitemporal data," *J. Appl. Remote Sens.* **5**(1), 053547 (2011), http://dx.doi.org/10.1117/1.3619838.
35. U.S. Department of Agriculture (USDA), "Trade and agriculture. What's at stake for California?," (2009).
36. USDA National Agricultural Statistics Service (NASS) and California Field Office, "California agricultural statistics 2011 crop year," (2012).

37. P. Canessa, S. Green, and D. Zoldoske, "Agricultural water use in California. A 2011 update," (2011).
38. E. F. Vermote, S. Y. Kotchenova, and J. P. Ray, "MODIS surface reflectance user's guide," pp. 1–40 (2011), http://modis-sr.ltdri.org.
39. R. E. Wolfe, D. P. Roy, and E. Vermote, "MODIS land data storage, gridding, and compositing methodology: level 2 Grid," *IEEE Trans. Geosci. Remote Sens.* **36**(4), 1324–1338 (1998), http://dx.doi.org/10.1109/36.701082.
40. R. G. Congalton and K. Green, *Accessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed., CRC/Taylor and Francis, Boca Raton (2009).
41. R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.* **37**(1), 35–46 (1991), http://dx.doi.org/10.1016/0034-4257(91)90048-B.
42. N. Ramankutty et al., "Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000," *Global Biogeochem. Cycles* **22**(1), GB1003 (2008), http://dx.doi.org/10.1029/2007GB002952.
43. R. S. De Fries et al., "Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers," *Int. J. Remote Sens.* **19**(16), 3141–3168 (1998), http://dx.doi.org/10.1080/014311698214235.
44. M. A. Friedl et al., "Global land cover mapping from MODIS: algorithms and early results," *Remote Sens. Environ.* **83**(1–2), 287–302 (2002), http://dx.doi.org/10.1016/S0034-4257(02)00078-0.
45. X. Xiao et al., "Landscape-scale characterization of cropland in China using Vegetation and landsat TM images," *Int. J. Remote Sens.* **23**(18), 3579–3594 (2002), http://dx.doi.org/10.1080/01431160110106069.
46. A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.* **25**(3), 295–309 (1988), http://dx.doi.org/10.1016/0034-4257(88)90106-X.
47. J. Qi et al., "A modified soil adjusted vegetation index," *Remote Sens. Environ.* **48**(2), 119–126 (1994), http://dx.doi.org/10.1016/0034-4257(94)90134-1.
48. G. Rondeaux, M. Steven, and F. Baret, "Optimization of soil-adjusted vegetation indices," *Remote Sens. Environ.* **55**(2), 95–107 (1996), http://dx.doi.org/10.1016/0034-4257(95)00186-7.
49. P. S. Thenkabail et al., "Assessing future risks to agricultural productivity, water resources and food security-how can remote sensing help?," *Photogramm. Eng. Remote Sens.* **78**(8), 773–782 (2012).

Zhuoting Wu is a GIS/remote sensing specialist interested in impacts of climate change on terrestrial ecosystems. She currently is a research ecologist in the Western Geographic Science Center of the U.S. Geological Survey, and works extensively on projects in GIS, passive and active remote sensing–based vegetation and land cover/land use mapping, in collaboration with multiple U.S. and international agencies.

Biographies of the other authors are not available.