Agricultural Systems 114 (2013) 95–105

Contents lists available at SciVerse ScienceDirect

Agricultural Systems

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

A 2010 map estimate of annually tilled cropland within the conterminous United States

David M. Johnson

United States Department of Agriculture, National Agricultural Statistics Service, United States

ARTICLE INFO

Article history:
Received 3 May 2012
Received in revised form 15 August 2012
Accepted 30 August 2012

Keywords:
Area
Classification
Cropland
Estimation
Mapping
Tillage

ABSTRACT

A ca. 2010, 30 m resolution map depicting annually tilled areas across the conterminous United States was developed. Input sources included four years, spanning 2008–2011, of annual national-level coverage Cropland Data Layer (CDL) land cover classifications as produced by the National Agricultural Statistics Service. Derived total land area under tillage from the aggregate CDL product equaled 112.8 million hectares (278.7 million acres). By comparison, the 2007 Census of Agriculture (CoA) produced an estimate of 122.9 million hectares, suggesting the map is under representing tilled area by 10.1 million hectares or 8.2%. Regression analysis using state-level summaries showed a strong, albeit biased, correlation (r-squared = 0.99) between the CDL derived tilled area and the CoA information. Notable outliers were North Dakota and Montana. Comparisons of the CDL tilled map were also made against the 2006 National Land Cover Dataset (NLCD) land cover product's Cultivated Crops category. Strong state-level regression agreement (r-squared = 0.98) was also found between the NLCD and the CDL acreages, but the NLCD estimated 8.5% more area than the CDL and thus closely matched that of the CoA. However, significant pixel level differences were found between the CDL and the NLCD. Nationally 5.6% of the maps were in disagreement as to whether cultivated or not, a large proportion considering around a seventh of the country's land area is tilled. States of Arkansas, Montana and Wisconsin had the largest absolute discrepancies between the NLCD and CDL. Accepting the CDL as reference showed a national level NLCD cropland commission error of 23.0% and omission error of 14.5%. Much of what is believed to be problematic in the NLCD could be explained by definitional issues having included alfalfa hay into their cultivated category for many areas. Ultimately, while it is likely that the CDL annually tilled area model is an underestimate of the true total, taken contextually in map form and adjusted for undercount bias it likely is the best available.

Published by Elsevier Ltd.

1. Introduction

1.1. Background

Understanding the area and extent of croplands is important for a variety of societal and environmental reasons. They include biodiversity (Landis et al., 2008; Tscharntke et al., 2005), pest and disease spread (Margosian et al., 2009; Werling et al., 2011), bioenergy management (Fargione et al., 2009; McDonald et al., 2009), watershed runoff and chemical monitoring (King et al., 2005; Meehan et al., 2011), wildlife management (Meehan et al., 2010), food security (Brown and Funk, 2008), water balance (Dappen et al., 2008) and weather (Raddatz, 2003; Steyaert and Knox, 2008). Also included in the list is the grand topic of climate change. Land use and land cover change, with the ramifications of carbon storage (Fargione et al., 2008; Jung et al., 2006; West et al., 2010) and greenhouse gas emissions (Searchinger et al., 2008), is not a fully understood driver for altered future global and United States (US) level climate scenarios (Feddema et al., 2005; Mahmood et al., 2010; Pielke et al., 2011). Because of agriculture's relatively large areal footprint, it likely plays a very large role in the climate system (Fall et al., 2009; Pielke et al., 2007). Impacts to the US are likely variable depending on location (Fall et al., 2010), and a shifting or more variable climate will likely alter what crops can be grown where (Schlenker and Roberts, 2009). All of the listed topics are interrelated and have complex feedback mechanisms (Foley et al., 2005).

In a more applied setting, it is also important to know how much land is utilized for resource allocation and economics (Rudel et al., 2009; You et al., 2009). If, for example, one is trying to plan for the transportation or utilization of crop products it is useful to have a detailed picture of the cultivated areas. Likewise, having a firm handle on the distribution of crop areas is helpful for natural disaster mitigation, assessment and response. And because growing areas are highly anthropogenic lands, they are likely more sensitive to change versus other land cover types. Within the US for

E-mail address: dave.johnson@nass.usda.gov

0308-521X/$ - see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.agsy.2012.08.004

example, farm policy or commodity pricing can have impacts on what gets planted where over the span of just a year.

The distribution and extent of croplands has been described and defined in many contexts. In tabular form, recent estimates are that 1381.2 million hectares (10.6%) of the global land area is arable by the Food and Agricultural Organization (FAO) of the United Nations (FAO, 2009). Within the conterminous US the FAO puts the number at 162.8 million hectares, or about 11.8% of the world's total arable land. Most countries independently estimate their amount of cropland through surveys or a census. Cropland areas of all types within the US are estimated at 164 million hectares by the US Department of Agriculture (USDA, 2009a). This represents about 18.0% of the total land cover in the conterminous US. Of that total, 125.3 million hectares (76.2%) is considered harvested cropland by the USDA.

1.2. Mapping cropland areas

Statistics alone may not shed detailed enough information on which areas of land cover are under cultivation or changing. Detailed maps, however, can also allow for extraction of the land cover statistics beyond what is typically provided at only national or sub-national levels. The development of global and regional crop map databases, primarily through remote sensing, is not new or unique. DIScover is a global one kilometer squared resolution Advanced Very High Resolution Radiometer (AVHRR) based land cover data set (Loveland et al., 2000) with categories for "croplands" and "croplands/natural vegetation mosaic." The University of Maryland also developed a one kilometer global product from AVHRR data, albeit with a different methodology (Hansen et al., 2000). Global Land Cover 2000, also one kilometer, was developed with SPOT 4 data (Bartholomé and Belward, 2005) with the land cover classes "cultivated and managed areas," "mosaic-Cropland/Tree Cover/Other natural vegetation" and "mosaic-Cropland/Shrub or Grass Cover."

More recent, occasionally updated, and at higher resolution, a Moderate Resolution Imaging Spectroradiometer (MODIS) based land cover product is available at 500 m (Friedl et al., 2010, 2002). It is named from the suite of MODIS derived products as MCD12Q1. It has the data categorized in several manners so as to be comparable to the one kilometer datasets mentioned previously. Embedded in MCD12Q1 are agricultural category names such as "croplands," "croplands/natural vegetation mosaic," "cereal crops" and "broad-leaf crops." Another global land cover classification product, and similar to MODIS, named GlobCover was developed from ca. 2005 Envisat Medium Resolution Imaging Spectrometer (MERIS) imagery (Arino et al., 2008). It is a finer 300 m resolution and has categories for crops including: "post-flooding or irrigated croplands," "rainfed croplands," "mosaic cropland (50–70%)/vegetation (grassland, shrubland, forest) (20–50%)" and "mosaic vegetation (grassland, shrubland, forest) (50–70%)/Cropland (20–50%)."

Similar map products have been developed that are just focused on US croplands. Chang et al. (2007) utilized MODIS 500 m data to map the distribution of corn and soybeans within the Corn Belt region of the US for 2002. Wardlow and Egbert (2008) created crop maps with 250 m data but focused instead on the US Great Plains. Also, Shao et al. (2010) used MODIS data but focused on the Great Lakes Basin region. All of these map products were shown to have good area agreement with official statistics from the USDA.

Even with all these sources of information the data still can lack the needed spatial detail, or accuracy for further analysis and often do not agree (Giri et al., 2005; Hansen and Reed, 2000; Herold et al., 2008). The timeliness or analyzed time period may not be useful either depending on the application. Improvement can be made by combining the map information with national surveys (Monfreda et al., 2008) but the importance of even higher resolution datasets was shown by Nelson and Robertson (2007), even when coarser scaled data are considered highly accurate.

Land cover products with finer resolve and information about agriculture do exist but are less common, particularly over large regions. Thirty meter grid cell size is arguably the de facto standard for these higher resolution products. Earliest derived was the Geocover LC dataset which has categories for "agriculture, general" and "agriculture, rice/paddy." It was produced ca. 1990 and encompassed the entire world (Cunningham et al., 2002). But, while ambitious at the time, it is dated now since two decades have passed from its analysis date.

The National Land Cover Database (NLCD) is a more recent and visible product but only over the US domain. It is also 30 m in resolution. There are currently three epochs; 2006, 2001 and 1992 (respectively, Xian et al., 2009; Homer et al., 2004; Vogelmann et al., 2001) and more are planned for the future. The NLCD is not solely focused on mapping cropland agriculture but has categories in its classification scheme that are directly related. The two most recent NLCD releases contained categories for "cultivated crops" while the older version contained categories for "row crops" and "small grains." The cultivated crops category specifically is defined as "areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class also includes all land being actively tilled."

Analysis of the NLCD agricultural classes against official USDA statistics has been performed in some contexts. Maxwell et al. (2008) assessed differences in the 2001 NLCD for the Upper Midwest portion of the US and Goslee (2011) did similar for the New England areas with the 2006 NLCD. Accuracy assessment results of all of the categories of the 2006 NLCD are yet forthcoming.

1.3. The Cropland Data Layer

Beyond all the datasets previously mentioned, the timeliest and arguably best crop tailored US land cover product is now generated annually by the USDA's National Agricultural Statistics Service (NASS). The dataset itself is coined the Cropland Data Layer (CDL). The CDL is typically a 30 m resolution raster-based grid spanning the conterminous US with focus on agricultural cover types. It goes into more category detail than just "cropland" or "cultivated crops" like found in the NLCD. Instead the CDL attempts to identify specific crop types like corn, soybeans, wheat, cotton, etc. Non-agricultural cover types are also documented but with a lesser degree of thematic detail. A general description of the 2009 product is given by Johnson and Mueller (2010) while a more in depth explanation of the CDL program is provided by Boryan et al. (2011). In short, datasets were created using Landsat Thematic Mapper (TM), or similar, multispectral satellite data in concert with "ground truth" reference information through a supervised classification methodology. There are currently US national level CDL coverages for each of the years spanning 2008–2011. But in total, NASS has been researching and developing these cropland tailored CDL land cover classifications over specific agriculturally intensive states for more than a decade.

The CDLs have several roles within NASS. The main use for the products has been to derive within season acreage estimates of the major commodity crops at state and county levels which help supplement data from ongoing and traditional NASS surveys. The CDLs serve as reference land cover information in roles such as developing land cover stratification for the NASS June-based and annually published Acreage survey (USDA, 2011). The CDLs help as a benchmark for defining cropland areas during times of natural disaster. And finally, the CDLs are used as ground truth for defining crop types in concert with coarser scaled data, such as MODIS, for crop progress monitoring and yield modeling. After the growing season is fully completed and NASS published estimates solidified, the CDLs are freely released for public consumption early the following year (Han et al., 2012; USDA, 2012).

The NASS land cover classifications usually have high thematic accuracies, particularly for the dominant commodity crops, but each category in the CDL still likely contains a certain amount of bias if used for estimating area from a simple "pixel counting" methodology. It is believed in general the CDLs have a tendency to underestimate crop areas modestly, but this undercount is not necessarily consistent year to year or region to region. Fortunately, any area bias can be compensated for through a "regression estimator" (Battese et al., 1988). This is accomplished by comparing area totals calculated at a sample of small areas within the CDL against estimates collected at the same areas by NASS field enumerators. The area relationship of what the CDL estimated versus the in situ estimates over the many samples form a linear relationship. Then the slope of the regression line providing the bias adjustment for what each pixel is truly worth in terms of area. Thus, acreage estimates NASS produces from the CDL are not simply a pure pixel area sum but rather the area of those pixels summed and then multiplied by regression parameters.

Earliest CDL efforts were research based and focused on the core Corn Belt states like Iowa and Illinois. But as the utility for the products has grown, the classification methodologies become more efficient, imagery datasets availability now at low cost or free, and computational power increased, the land cover product became national in scope. The year 2009 marked the first national coverage that was completed during the same growing season. Since then CDLs have been produced and released nationally for the 2010 and 2011 growing season. The year 2008 was not complete nationally initially but retrospective analysis was performed for the states not originally undertaken. As a result, a full conterminous US 2008 CDL now also exists.

Thus, there are four years of complete CDL coverage for the conterminous US spanning the years 2008–2011. The native pixel resolution of the 2010 and 2011 CDLs is 30 m with a map projection of Albers Conic Equal-area. The 2009 product's resolution is 56 m which coincides with the at nadir sample size of the Resourcesat-1 Advanced Wide Field Sensor (AWiFS) imagery that was used primarily by the USDA during that period. 2008 is mostly 56 m resolution, again due to the utilization of AWiFS. However, the several states that were classified retrospectively utilized Landsat as the primary data set and thus have native resolution of 30 m. Those states are California, Florida, Idaho, Washington, Oregon, Montana, Michigan and all those in the New England region.

Accuracy assessments of the crop category aspects of the CDLs were undertaken at the state-level and available in the metadata for each. Full discussion of the accuracies is beyond the scope here. In summary though, state-level accuracies at the pixel-level for the individual tilled crop type categories typically ranged from about 70.0% to 95.0%. For year 2011 the non-weighted, by state size, crop accuracy average was 85.3%. Year 2010's accuracies were a percentage point lower at 84.3%. The averages for 2009 and 2008 were notably lower with overall crop accuracies around 80.0% and 76.9%, respectively.

States with the highest accuracies tended to be those in the Corn Belt region. This is likely a function of the regimented nature of the cropping practice, availability of the most complete ground truth, and premium by NASS placed on acquiring the best imagery in that area since it makes up the bulk of US commodity production. CDL accuracies have steadily improved through the years probably due to the ability to analyze larger volumes of time-series imagery, more complete ground truth, and processing experience. Additionally, the older 2009 and 2008 products were primarily reliant on the lower resolution AWiFS imagery and thus a coarser 56 m versus 30 m grid cell size was used for most of those classifications. While it is believed the larger AWiFS pixels do not overly impact crop statistics because most fields in the US are relatively big in area and homogenous to begin with, there is still an expected accuracy loss attributable to the decreased ability to map detailed landscape features like complex field edges.

1.4. US cropland area non-map statistics

Other sources of US crop area statistics are important even though they do not have a direct map component. The NASS Census of Agriculture (CoA) provides the most complete and detailed inventory of agricultural land use available for the US. The CoA is performed every five years with the most recent undertaking during the 2007 calendar year (USDA, 2009a). CoA statistics are primarily aggregated at county, state, and national levels. Additionally, tabulation of the data was performed at watershed and congressional district levels for 2007. Much of the CoA information is also provided in atlas form, usually through simple choropleth maps, but with no additional detail beyond county level. Agricultural and use statistics pertaining to the area tilled or cultivated are not published directly by the CoA. They can however be calculated by reconciling the land use tables. Case in point, total cropland is published but it includes by definition areas used for pasture and grazing, forage crops, orchards, nursery stock, and short rotation woody crops. Those subcategories have tables of their own and thus can be subtracted from the overall totals however. The CoA also details in Appendix A (USDA, 2009b) overall measures of quality and reliability. Because it is a census and not a survey much of the potential error is based on under-coverage or non-response. Farm-level adjustments and relative errors are listed but do not translate directly to potential errors in the land use information.

The other main NASS source of cropland utilization information, and which is produced more timely and annually, is known as Acreage (USDA, 2011). It is published the last business day of every June and is the summation of results from the NASS area based June Acreage Survey and the list based June Agricultural Survey. They are both undertaken the first half of that same month. The survey sampling methodology is large in that about 2.9 million hectares of land, with unit sizes of about 260 ha each, is visited and enumerated each year. The survey is tailored toward the US and state-level area statistics of the large commodity crops. It also does not include information at geographic levels more detailed than US states. Furthermore, the statistics are not mutually exclusive for a given parcel of land over a growing season. This means some areas with two or more crops per year on the same piece of ground get double counted in the land use total. So for these several reasons combined, it is not perfectly suited for spatially detailed or unbiased total crop area estimates.

2. Methodology

2.1. Defining annually tilled

A clear definition of "annually tilled" was first derived. For this work tilled included all areas that traditionally have annually seeded crops with manipulation, either physically or chemically, of the soil in between or during growing seasons. This included all of the primary field crops like corn, soybeans, wheat, rice, and cotton along with vegetable crops and non-tree fruit crops. Perennial crops such as hay were excluded from the definition. Orchard and tree crops were also not included. Ultimately, usage of the term tilled in this paper's context is meant to be more exclusive than the nearly synonymous and commonly used word "cultivated." Cultivated by some definitions would also include tree, perennial crops, and/or livestock areas.

Potentially confusing the definition of tilled, some farmers practice "conservation tillage" or "no-till" soil management techniques. These are fields, where the soil is not turned by plow and new seeding is planted directly into the previous crop's residue. However, these areas were indeed included in the definition of tilled being undertaken here. And confusing the annual portion of the definition, areas that typically sit idle or "summer fallow" for a year because crops are planted only biennially were also included. The most common example of this is for wheat crops in the dry areas of the western Great Plains. Likewise, fields could have had more than one planted crop per year were also included.

A couple of CDL crop categories demanded special attention as to their tilled status or not. Alfalfa is found throughout the US and often rotated with field crops, particularly corn, after several years of residing perennially. It was not included in the tilled category for this analysis although is some context it could be argued it should. NASS also generalizes alfalfa as a hay crop and that is how it was ultimately considered for this analysis. Conversely, the CDL category seed/sod grass, most abundant in the Pacific Northwest region, was included in the tilled definition. These grass areas are heavily managed annually with similarities more in line with field crops than hay. Furthermore, seed and sod grasses are not categorized as field or hay crops by NASS. They are a special category altogether.

2.2. Map derivation of tilled areas

With the tilled definition in place, four years of the national level CDLs spanning the years 2008–2011 were assimilated to derive a best of ca. 2010 annually tilled area classification. The integration of multiple years provided the opportunity to develop a more refined product than what a single years' CDL could provide. This is particularly true given every years' CDL, like any thematic map, has a certain amount of misclassification or noise.

A simple rule-set was constructed to build the tilled map from the four input years. First, each CDL year was categorically recoded to "tilled" or "not tilled" based on the appropriate raw input categories (Table 1). Then, all four years were "overlaid" and assessed as for tilled status at each pixel location. If two or more of the four total input years were in the tilled category then the output classification was called tilled.

One exception was not allowed – if only the pixels in years 2008 and 2009 were deemed tilled then the output pixel was not called tilled. In other words, minimally either 2010 or 2011 had to have been called tilled for the output to reflect it too. This modification was implemented for two reasons. One, all of the 2009 and most of the 2008 CDLs were coarser in spatial resolution and with less thematic accuracy than the 2010 and 2011 products. Thus, there was somewhat less confidence in them which meant they were more likely to propagate error. The second reason for the less weight applied to 2008 and 2009 was the overall goal to develop something closest to a 2010 era product. Thus it was also desirable to give the older 2009 and 2008 products less weight in the rule voting decision.

Two special categories, alfalfa and orchards, were also labeled in the output for later comparison and reconciliation against other data sets, particularly the most similar NLCD. Alfalfa was carried as unique category for two reasons. One, in some contexts it could be considered a tilled crop because usually it gets rotated with other crops, particularly corn, after a few years. Secondly, prior observation of the NLCD by NASS has shown alfalfa suspect of often being included in its NLCD cultivated crops category instead of pasture/hay, which is more appropriate by the NLCD's own definition. Orchards were carried as a second a unique category in order to better compare to the NLCD as its cultivated croplands category also includes orchards by NLCD definition. Ultimately, the voting rule methodology for alfalfa and orchards were implemented in the same manner as the tilled crops.

Finally, after the voting rules derived the new single output layer, a minimum mapping unit of five pixels (about 0.4 ha or one acre) was applied to the entire output image to reduce some of the "speckling." Single pixels or very small patches of contiguous pixels had a high probability of being spatial noise since crop field sizes tend to be relatively large in area. Moreover, groups of pixels were not considered conjoined if they only touched on the corners. This helped reduced "stringy" or linear features which are also unlikely to be crops areas.

The analysis was performed independently for each state. There were two reasons for this. One, because there was no way to best harmonize the 2008 CDL into a single mosaic without compromise given the mix of 30 m and 56 m resolutions. And two, application of the minimum mapping unit is computational intensive and was not possible to run over the entire US at once. Output projection was the same as common to all of the input CDLs- Albers equal-area conic. The output grid was set to 30 m resolution and nested the same as the 2011 and 2010 input CDLs. The state analysis boundaries were defined on the Census-based Shapefiles "dtl_st.shp" distributed by Environmental Systems Research Institute (ESRI) via their ArcGIS software suite. All processing steps were performed in Erdas Imagine 2011.

Table 1
CDL numeric codes with category names placed into tilled classification output.

1. Corn	27. Rye	48. Watermelons	213. Honeydew melons
2. Cotton	28. Oats	49. Onions	214. Broccoli
3. Rice	29. Millet	50. Cucumbers	216. Peppers
4. Sorghum	30. Speltz	51. Chick Peas	219. Greens
5. Soybeans	31. Canola	52. Lentils	221. Strawberries
6. Sunflowers	32. Flaxseed	53. Peas	222. Squash
10. Peanuts	33. Safflower	54. Tomatoes	227. Lettuce
11. Tobacco	34. Rapeseed	55. Caneberries	229. Pumpkins
12. Sweet corn	35. Mustard	56. Hops	242. Blueberries
13. Popcorn/ornamental corn	38. Camelina	57. Herbs	243. Cabbage
	41. Sugarbeets	59. Sod/seed grass	244. Cauliflower
14. Mint	42. Dry beans	61. Fallow/idle cropland	245. Celery
21. Barley	43. Potatoes	205. Triticale	246. Radishes
22. Durum wheat	44. Other crops	206. Carrots	247. Turnips
23. Spring wheat	45. Sugarcane	207. Asparagus	248. Eggplants
24. Winter wheat	46. Sweet potatoes	208. Garlic	249. Gourds
25. Other small grains	47. Miscellaneous vegetables and fruits	209. Cantaloupes	250. Cranberries

[a] Not shown were the also included "double cropped" classes: 26, 225, 226, 230–241, and 254.

3. Results

3.1. Simple metrics

Direct pixel counting of the ca. 2010 annually tilled classification resulted in a total area estimate of 112.8 million hectares (278.7 million acres) within the conterminous US. This represents about 14.7% of the total land area (i.e. not including water) available. An overview map showing the distribution of tilled areas is shown in Fig. 1. Not surprisingly, the majority of the tilled area is found in the mid-section of the US in the Corn Belt and Great Plains regions. Significant density of tilled area is also found in the California Central Valley, The Delta (the Mississippi Alluvial Plain) and the Intermountain region of the Northwest.

A complete list of state level area totals are provided in Table 2. The top five states for annually tilled area are Iowa, Kansas, Illinois, Texas and North Dakota. Two of those states, Iowa and Illinois, represent areas for which more than half of the land is utilized for tilled cropping. The five states with the least amount of tilled area are Rhode Island, New Hampshire, Connecticut, Nevada and West Virginia. The first three are no surprise since their land area is very small to begin with. Nevada is very large though and its tilled area is the smallest percentage of any across the US at only 0.1%. West Virginia and New Hampshire both also resulted in low areas of tilled with less than 1.0% of the total land area. The District of Columbia had no tilled cropland.

3.2. Comparisons to Census of Agriculture statistics

The 2007 CoA produced an estimate of 122.9 million hectares (303.6 million acres) of tilled land after subtracting pasture and grazing, forage crops, orchards, nursery stock, and short rotation woody crops from the overall cropland category. Using the CoA as truth, the suggestion is the CDL derived tilled map is under representing the actual area by 10.1 million hectares (24.9 million acres), or about 8.2%. A scatter plot showing the relationship between the CDL and CoA tilled areas tallied at the state-level is depicted in Fig. 2. The linear relationship between the two variables overlaid is also shown and with a very strong correlation with an *r*-squared of almost 0.99. The bias is graphically documented with a least-squares regression line overlaid the chart. The CDL tilled likely underestimating the CoA tilled as the slope of the line does not equal one. The calculated relationship is CoA tilled = 1.088 $*$ CDL tilled. The root-mean-square error (RMSE) is 0.5 million hectares.

The most notable outliers undercounting are Montana (MT) and North Dakota (ND). They are beneath the CoA counts by about 28.0% and 18.0%, respectively. Texas (TX) has the biggest departure from the regression line on the other side. But it falls almost perfectly on the 1-to-1 line so there is ultimately no suggestion of bias for it. But, it is probably an overestimate of tilled area given the relationship to the rest of the points. For states with small tilled areas it is difficult to assess the relative amount of difference from Figs. 1 and 2 since they have little area to begin with. Delaware, a relatively intensive cropland state, would be the biggest outlier percentage area wise though with suggestion that CDL-based tilled estimate statistics are over counting the true tilled area percentage by about 20.0%.

3.3. Comparisons to NLCD statistics

The total conterminous US area from the 2006 NLCD "cultivated crops," category number 82, is 124.3 million hectares (307.3 million acres) obtained via direct pixel counting. The sum was derived by using the same boundary extent file as for the CDL analysis. The NLCD numeric is not directly comparable to the CDL estimate because it includes orchards in its definition. Thus, to make the comparison valid the tabulation of the special orchards CDL derived class was added to the CDL tilled area summaries. The orchard adjusted CDL total is 114.5 million hectares (283.0 million acres) which is about 1.5% larger than without. The resulting difference between it and the NLCD is 9.8 million hectares, or about 7.9%, and similar to what was found differentiating the CDL tilled from the CoA tilled.

The state-level regression comparisons are shown in Fig. 3. Again, the CDL estimates here include orchards in order to compensate for the NLCD definition which includes it. Henceforth the term CDL cultivated will be used to reference it. The relationship (NLCD cultivated = 1.055 $*$ CDL cultivated) is very strong with an *r*-squared over 0.98. The RMSE is 0.4 million hectares. The largest outlier differences, where the CDL cultivated area estimates are less than the NLCD are for the states of Wisconsin (WI) and Montana (MT). Conversely, Arkansas (AR) appears it may be overestimated in the CDL cropland analysis. If the statistics are normalized for land area, Wisconsin is still an outlier, underestimating cultivated area, and both Arkansas and Delaware overestimating.

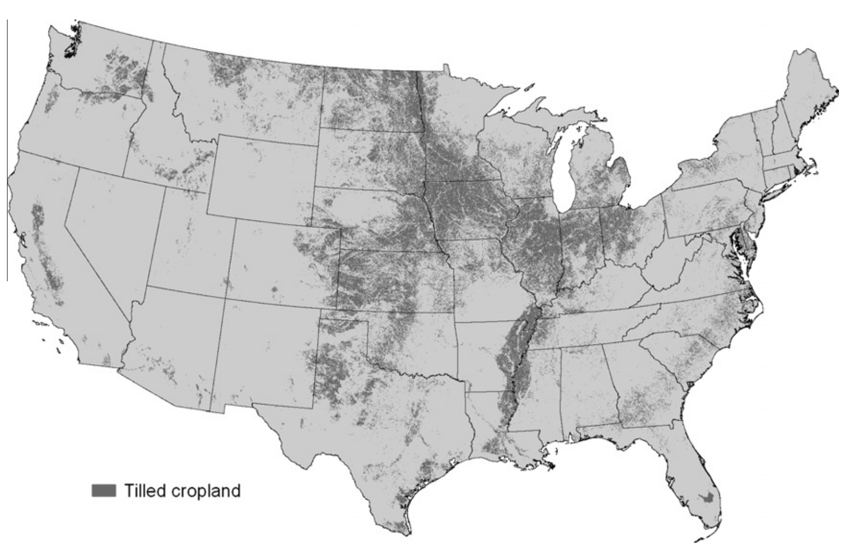

Fig. 1. Overview map of CDL derived tilled cropland areas in the conterminous US.

Table 2
Estimates of tilled area and% tilled by US state from the CDL analysis.

State	Tilled area (ha)	Land% tilled
Alabama	522,813	4.0
Arizona	537,522	1.8
Arkansas	2,733,620	20.3
California	1,812,516	4.5
Colorado	2,755,078	10.3
Connecticut	21,769	1.7
Delaware	197,654	39.2
Florida	516,279	3.7
Georgia	1,294,885	8.6
Idaho	1,380,605	6.4
Illinois	8,679,430	60.5
Indiana	4,523,148	48.8
Iowa	9,296,238	64.4
Kansas	9,147,169	43.4
Kentucky	1,026,153	10.0
Louisiana	1,809,282	16.4
Maine	101,226	1.3
Maryland	501,222	20.0
Massachusetts	30,156	1.5
Michigan	2,437,007	16.6
Minnesota	7,086,247	34.5
Mississippi	1,634,770	13.5
Missouri	3,410,465	19.2
Montana	3,976,685	10.5
Nebraska	6,794,950	34.2
Nevada	21,761	0.1
New Hampshire	7,764	0.3
New Jersey	180,933	9.5
New Mexico	544,980	1.7
New York	792,801	6.5
North Carolina	1,658,771	13.2
North Dakota	8,033,671	45.4
Ohio	3,466,048	32.8
Oklahoma	2,940,988	16.6
Oregon	1,014,968	4.1
Pennsylvania	937,536	8.1
Rhode Island	3,025	1.1
South Carolina	738,112	9.5
South Dakota	5,484,602	28.1
Tennessee	1,036,037	9.7
Texas	8,272,463	12.2
Utah	242,765	1.1
Vermont	45,801	1.9
Virginia	526,695	5.2
Washington	1,901,730	11.0
West Virginia	25,391	0.4
Wisconsin	2,443,661	17.4
Wyoming	255,258	1.0
Conterminous US	112,802,649	14.7

3.4. Comparison to NLCD map

Aggregated total statistics do not necessarily tell the whole story. A region could have very similar totals but may not match geographically at the local, field, or pixel level. Overlay analysis was thus performed to count the pixel level agreement between the cultivated areas in NLCD and CDL. Direct comparison of the two within a geographic information system was straightforward and perfectly aligned since NLCD and CDL tilled share the same map projection, cell size and reference grid origin. For all pixels, US level agreement for non-cultivated pixels (agriculture or otherwise) was 81.7%. For the cultivated cover type the agreement was 12.7%. In the NLCD 3.4% of the pixels were labeled cultivated that were not in the CDL. And finally, the remaining 2.1% were cultivated in the CDL but not in the NLCD. The total does not exactly equal 100.0% due to rounding. If one were to assume the CDL cultivated product to be the truth or reference, then the commission error in the NLCD cultivated crops is 23.0% and omission error 14.5%.

Pixel agreement result percentages by state are shown graphically in Fig. 4. Most states have total pixel level agreements above 90.0%. Largest disagreements are in the states of Wisconsin and Delaware at near 17.0%. Wisconsin is particularly large discrepancy in being non-cultivated in the CDL and cultivated in the NLCD. Delaware is the opposite with a large CDL component that is cultivated, where the NLCD is not. Less severe and more balanced differences that are still notable are Indiana, Iowa, Kansas, Maryland, North Dakota, Ohio, and South Dakota. All are relatively heavy cropland states and have disagreements a bit above 10.0%. Granted, the statistics can be misleading because by random chance it is easiest to have agreement in states with little percentage of crop areas.

If one of the classifications is taken as truth, then it is possible to compare the disagreement scenarios to obtain assessment of omission and commission errors and a sense of bias between them. For example, Kansas which is one of the more intensive cropland states, shows 5.1% of the pixels are labeled as cultivated in the NLCD and not cultivated in the CDL, while 4.9% of the pixels are the opposite. Thus, there is a total disagreement 10.0%. But because the errors are nearly equal but contrasting there is little overall bias. This is a prime example, where the state totals do not tell the whole story. The state with the biggest difference between the two error metrics is Wisconsin. 13.4% of the pixels there are categorized as cultivated in the NLCD and non-cultivated in the CDL. The opposite scenario is 4.0%. Thus, their errors do not cancel and it suggests the NLCD opinion on cultivated area is much larger.

Conterminous US wide map differences between the NLCD and CDL cultivated are shown in Fig. 5. Notable regional discrepancies were found across the mid-Atlantic's Delmarva Peninsula, south-central Colorado and central Wisconsin. The latter area is shown in detail in an example map subset in Fig. 6c. Many of the map differences appearing visually can be attributed to the inclusion of areas of alfalfa hay into cultivated crops within the NLCD product. Cross checking against the CDL derived alfalfa category areas helps confirm this. While by some definitions alfalfa could be considered a cultivated crop, by strict interpretation of the NLCD definition it best fits into pasture/hay and thus is in error. For clarification, pasture/hay is defined by the NLCD as "areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation." Also showing lots of disagreement against the NLCD are areas of "summer fallow" within the CDL cultivated map. Summer fallow are fields that are kept out of production during a regular growing season and thus may only be tilled every other year. But, many of these summer fallow areas appear to be incorrectly as pasture/hay in the NLCD.

4. Discussion

Tilled areas results from the CDL analysis aligned closely with the CoA at the state level but showed a general undercount bias. This is consistent with what is perceived as underestimation by the CDL crop categories independently. By using the CoA as reference though it allows for the tuning of these CDL tilled area statistics through regression estimation. There were a total of 112.8 million hectares (out of 765.1 million hectares of total land area) found to be annually tilled in the entire conterminous US. Multiplying them by the CDL versus CoA regression slope of 1.0879 yields an adjusted area of 122.7 million hectares. This is just 0.1 million hectares below the CoA total. One can apply the 1.0879 multiplier to all pixels across a given study areas to come up with what are likely more refined area totals, regardless of geographic extent.

States appearing as outliers in the regression analysis between the CDL and CoA make it suspect that there is an error in one,

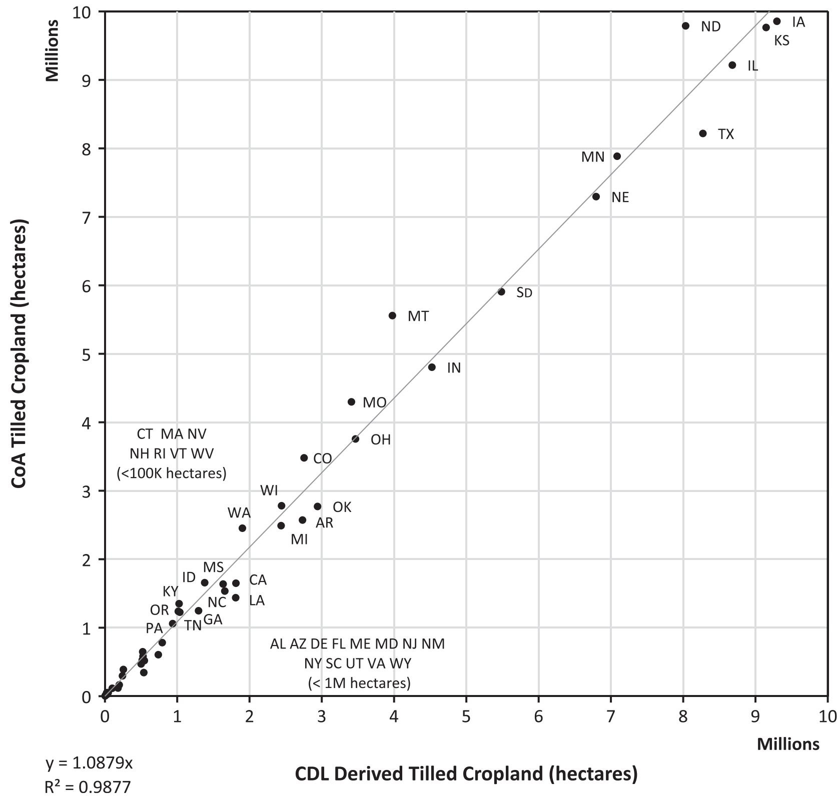

Fig. 2. Area estimates of CDL derived tilled cropland versus CoA tilled cropland.

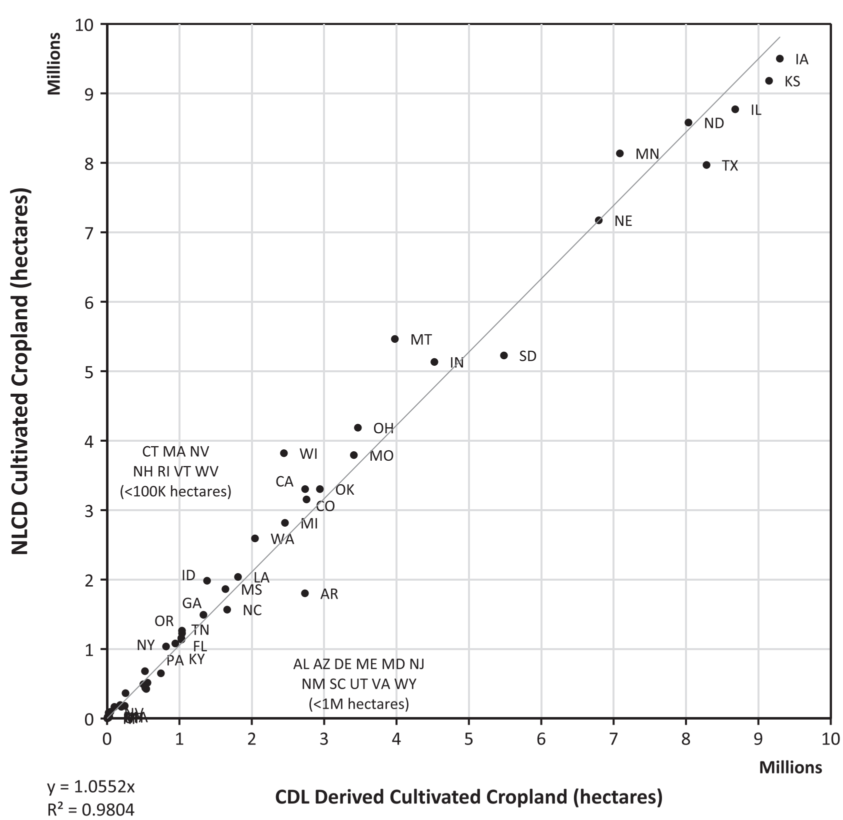

Fig. 3. Area estimates of CDL derived cultivated cropland versus NLCD cultivated cropland.

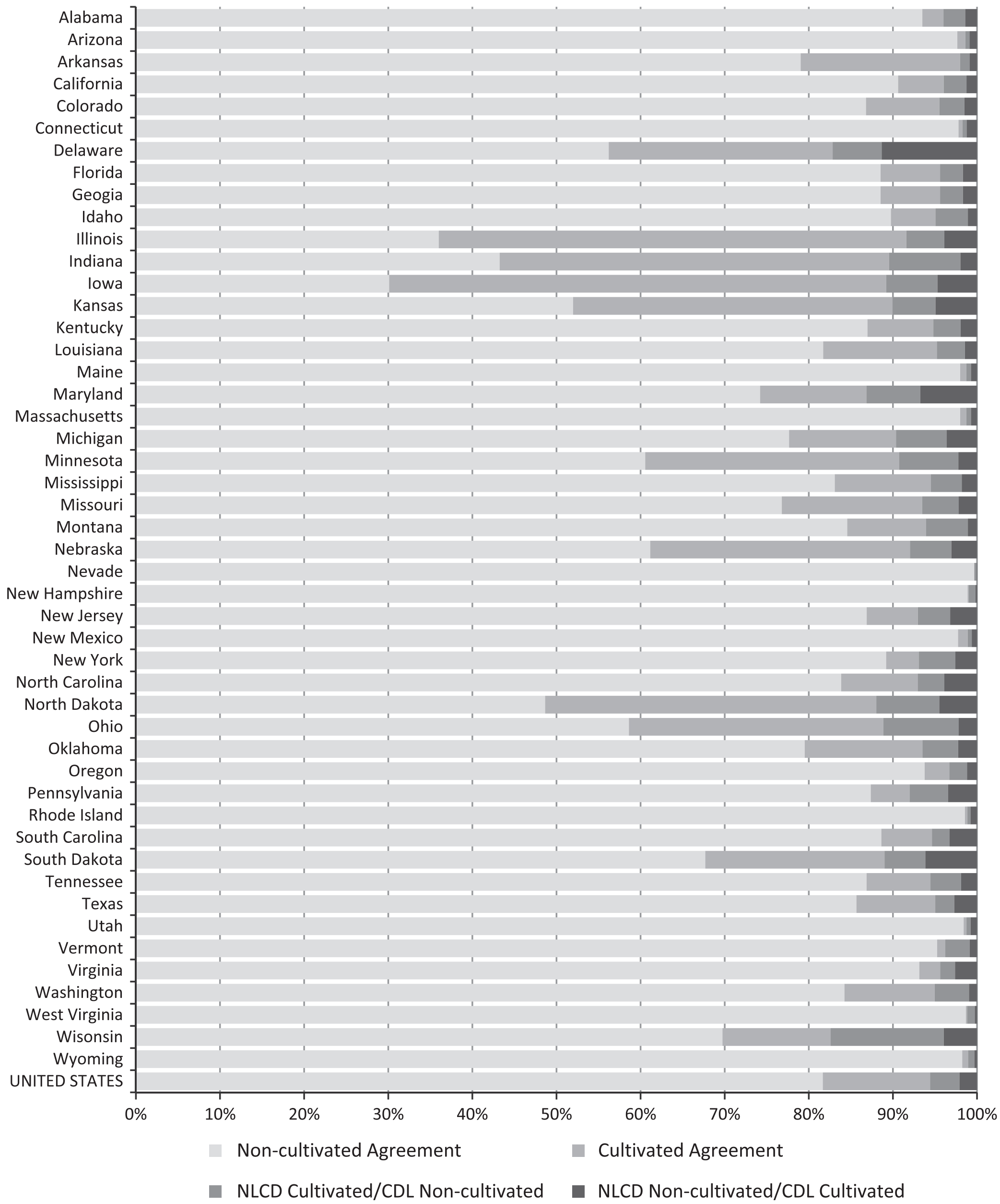

Fig. 4. State-level pixel agreement between cultivated in the CDL derived versus NLCD.

the other or both. It is probably more likely in the CDL analysis but it could be interpreted as a weakness in the CoA data as well. The states with the largest differences would be the first to investigate if trying to gauge the reliability of either dataset. Montana, for example, has a large absolute difference with the CDL tilled map about 1.6 million hectares under the CoA area estimate, and of all the states it falls farthest from the regression line. The discrepancy may lie in the fact there is a lot of dryland type farming in which many fields may sit for several years without being used. Thus, the CoA may capture them as cropland, as that is how they were reported by the farmer, but nothing was actually actively grown on them during the 2008–2011 CDL years. A similar situation is probably also occurring in neighboring North Dakota which actually has the largest difference (nearly 1.8 million hectares) between the tilled area analysis and the CoA.

Crop area estimate differences could be compounded by the non-overlap of the years used in the analysis between the CDL information and both the CoA and NLCD information. In the former case there is a three year difference (2007–2010) and the later four years (2006–2010). This is always a caveat between dataset inter-comparisons when the timing does not coincide. And, the larger the time differences between the analysis the larger the uncertainty. While it is believed there were not significant changes between the areas tilled from one year to the next, there were certainly at least some areas that went into or out of crop production. This is particularly true since commodity prices have risen significantly in the last few years and thus likely more total land had gone into production as the demand for grain has increased. The forthcoming 2012 CoA will best establish what direction land change has occurred after 2007. To gain prior insight, change analysis of the CDLs from 2008 and 2011 could be undertaken and potentially help quantify the differences. But, careful attention to the error rates of the different CDLs needs to be incorporated so that any change found is truly reflective of what is occurring on the ground and not just the compounding of inherent noise or bias.

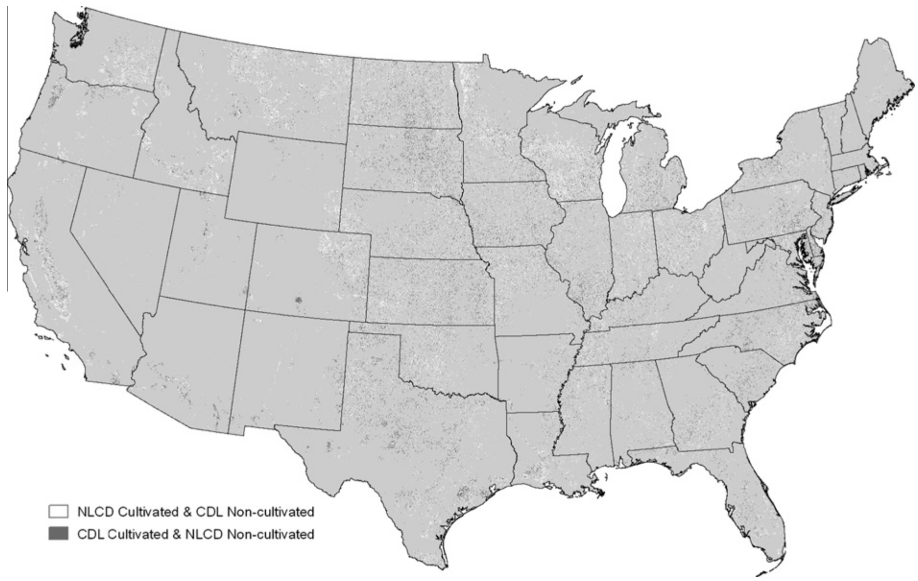

Fig. 5. Map differences between CDL cultivated cropland and NLCD cultivated cropland.

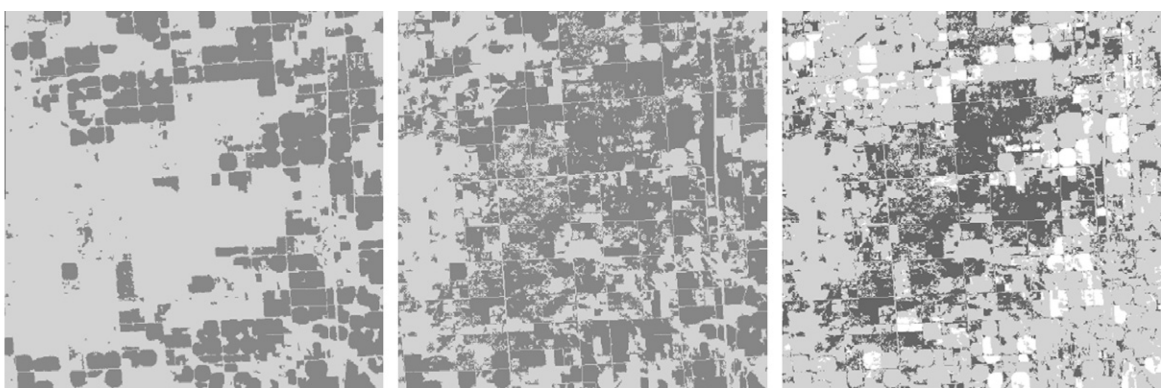

Fig. 6. Map subset example comparison of CDL and NLCD cultivated areas (central Wisconsin): (a) CDL derived cultivated (dark gray), (b) NLCD cultivated (dark gray) and (c) difference image (see legend in Fig. 5).

When comparing the NLCD versus the CDL cultivated areas it is reasonable to ask which is more correct. This is difficult to firmly answer without robust ground reference information about tilled versus non-tilled cover types from the concurrent year to validate either product. The CDL crop category accuracies vary by crop type but are generally in the 85–95% correct range for the major field crops which dominate what was ultimately being labeled as tilled. The combining of the individual crop categories should be improving those accuracies further. The assimilation of four years of CDLs to derive a single product should even more improve the accuracy over information from a single year product like the 2006 NLCD. And finally, areas found here with large discrepancies were investigated and believed to be more likely in error in the NLCD than the CDL. Thus, it is believed that in terms of cultivated areas the area estimates and maps from the combined CDLs are more robust than what the NLCD provides.

Many of the egregious differences between the CDL and NLCD come from the initial categorization of alfalfa and fallow crop areas. From the CDL perspective, which is also likely applicable to the NLCD, these are two of the most difficult classes to map. For alfalfa, and other hays constituting forage crops, the CDLs only map to an estimated accuracy of roughly 50% and are often highly confused with pasture and grasslands. In many regions this is no better than a simple educated guess. For the fallow class, it is really a usage category and thus may not be portraying a consistent cover type or age. For example, in some areas fallow may represent bare soil, for a single season, while in other areas it may be grass or trees if the ground has not been utilized for crops for years. This variability makes the CDL's classification process for fallow a deeper challenge since it is based both on spectral reflectance and temporal properties in its decision process. Furthermore, even if starting with highly accurate land cover maps the harmonization of different category definitions can be difficult and there is no true standard. It is generally believed here that both alfalfa and summer fallow are better represented in the CDL but some of the differences could be perceived as definitional, particularly within the NLCD. And, it is also acknowledged that both alfalfa and fallow areas could be arguably both be suited or not suited as cultivated crop categories given a particular context.

Only state level regression estimates were given here. More robust analysis could be undertaken with the same methodology but utilizing data aggregated to the county level. There are over 3000 counties nationally which would greatly expand the number of data points. Analysis could be done nationally as a whole or regionally to fine tune regression estimates and better ascertain localized differences between the tilled CDL area statistics and those from the CoA or NLCD.

The multi-year CDL aggregation methodology chosen may not be the best possible or could potentially be refined further. Less weight in the voting rule was given to years 2008 and 2009 since they had lower thematic accuracies to begin with and were usually based on the 56 m AWiFS resolution. That decision was ultimately determined subjectively, after some trial and error, with not only a goal of accuracy but also of simplicity and consistency across the entire US. It could be argued that the data from 2008 and 2009 should not have been included at all, voting rules made different and more complex, or a more sophisticated spatial smoothing routine employed. Furthermore, it is reasonable to imagine that different regions may work better with different techniques so that a single universal approach may not be a best fit for all areas.

The emphasis on this work was to document tilled versus non-tilled areas. The analysis could easily be modified to target specific crops or crop types. The weight or usage rules of the different years could also be altered to better tailor the output results. Derived data layers examples like "high probability corn areas," "C5 plant areas" or "small grain crops" could be easily constructed. Furthermore, analysis of crop rotation patterns could be undertaken now that a time series of annual crop specific land cover information has become available.

The analysis presented has been broad scaled. Admittedly, the national and state-level statistics presented from the CDL probably do not add much value to what is found in the CoA other than add some caution in cases found to be in relatively wide disagreement. However, the real appeal of the this type of derived dataset is when statistics or land cover spatial analysis is needed at local or even field level geographic scales. In those cases only a high quality land cover product, like the annually tilled CDL layers accomplishes the goal.

5. Conclusion

Area and map estimates of annually tilled croplands have been presented for the conterminous US. The ca. 2010 product was derived by assimilating four years of CDL data to create the best and most consistent national level map possible. These CDL data have provided a new era of spatially detailed and timely land cover of croplands available across the US. The ultimate goal of this work was to derive the very best model of land considered to be under annual tillage practice in the US and, alongside, understand the result's errors or biases.

It was shown the CDL derived annually tilled area to be about 112.8 million hectares nationally. This underestimates the true area by about 8.2% when using the CoA as reference. Regression analysis at the state level showed the underestimate bias to be about 8.8%. Either way, some added uncertainty of the difference exists since the CoA was based on the 2007 crop year and the tilled map more closely represents 2010. Thus, there may have been some true land cover change during the same time confounding the comparison of the results. States with the most notable differences from the CoA were Montana, North Dakota and Texas. Delaware was also an outlier when looked at from a percentage difference standpoint.

An evaluation of the NLCD cultivated cropland category was also undertaken. Area total statistics for cultivated cropland in the NLCD closely matches that of the CoA at national and state levels. However, map differences, sometimes large, between the NLCD and the CDL tilled were found which put into question the accuracy of the NLCD cultivated layer. Of particular issue is the inclusion of areas likely to be alfalfa into the NLCD cultivated cropland areas which by definition should not include them. States showing the largest differences by cultivated area percentage were Arkansas, Delaware and Wisconsin.

Detailed and timely land cover information is usually difficult to find and often lacking in quality, resolution, or scope. Datasets like the annually updated 30 m CDL are changing this though. And while it is likely that the CDL tilled model derived here is an underestimate of the true area total, taken contextually in map form and adjusted for bias it should be considered state of the art. It is particularly suited to regional or local land cover analysis of croplands, where high resolution maps or area statistics just do not exist.

Acknowledgements

Special thanks to colleagues whom provided suggestions and editing help. Particularly, they were Jeff Bailey, Claire Boryan, Rick Mueller, Bob Seffrin and Patrick Willis. Also thanks to the peer reviewers each of whom added useful and thoughtful improvements.

References

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., Weber, J.L., 2008. GLOBCOVER: the most detailed portrait of the earth. Eur. Space Agency Bull. 136, 24–31.

Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26 (9), 1959–1977.

Battese, G.E., Harter, R.M., Fuller, W.A., 1988. An error-components model for prediction of county crop areas using survey and satellite data. J. Am. Stat. Assoc. 83 (401), 28–36.

Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. Geocarto Int. 26 (5), 341–358.

Brown, M.E., Funk, C.C., 2008. Food security under climate change. Science 319 (5863), 580–581.

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. Agronomy J. 99 (6), 1654–1664.

Cunningham, D.J., Melican, J.E., Wemmelmann, E., Jones, T.B., 2002. GeoCover LC – A Moderate Resolution Global Land Cover Database, ESRI. <http://proceedings.esri.com/library/userconf/proc02/pap0811/p0811.htm> (accessed 13.08.12).

Dappen, P.R., Ratcliffe, I.C., Robbins, C.R., Merchant, J.W., 2008. Mapping agricultural land cover for hydrologic modeling in the Platter River watershed of Nebraska. Great Plains Res. 18, 39–52.

Fall, S., Diffenbaugh, N.S., Niyogi, D., Pielke Sr., R.A., Rochon, G., 2010. Temperature and equivalent temperature over the United States (1975–2005). Int. J. Climatol. 30, 2045–2054.

Fall, S., Niyogi, D., Gluhovsky, A., Pielke Sr., R.A., Kalnay, E., Rochon, G., 2009. Impacts of land use land cover on temperature trends over the continental United States: assessment using the North American Regional Reanalysis. Int. J. Climatol. 30, 1980–1993.

Fargione, J.E., Cooper, T.R., Flaspohler, D.J., Hill, J., Lehman, C., McCoy, T., McLeod, S., Nelson, E.J., Oberhauser, K.S., Tilman, D., 2009. Bioenergy and wildlife: threats and opportunities for grassland conservation. Bioscience 59 (9), 767–777.

Fargione, J., Hill, J., Tillman, D., Polasky, S., Hawthorne, P., 2008. Land clearing and the biofuel carbon debt. Science 319, 1235–1237.

Feddema, J.J., Oleson, K.W., Bonan, G.B., Mearns, L.O., Buja, L.E., Meehl, G.A., Washington, W.M., 2005. The importance of land-cover change in simulating future climates. Science 310, 1674–1678.

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global consequences of land use. Science 309, 570–574.

Food and Agriculture Organization (FAO) of the United Nations, 2009. FAOSTAT/Resources/Land/Arable, 2009. <http://faostat.fao.org/site/377/DesktopDefault.aspx?PageID=377#anc> (accessed 13.08.12).

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley., A., Huang, X., 2010. MODIS collection five global land cover: algorithm refinements and characterization of new dataset. Remote Sens. Environ. 114 (1), 168–182.

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 (1–2), 287–302.

Giri, C., Zhu, Z., Reed, B., 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. Remote Sens. Environ. 94 (1), 123–132.

Goslee, S.C., 2011. National land-cover data and census of agriculture estimates of agricultural-use area differ in the Northeastern United States. Photogramm. Eng. Remote Sens. 77 (2), 141–147.

Han, W., Yang, Z., Di, L., Mueller, R., 2012. CropScape: a web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. Comput. Electron. Agric. 84, 111–123.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int. J. Remote Sens. 21 (6 and 7), 1331–1364.

Hansen, M.C., Reed, B., 2000. A comparison of the IGBP discover and University of Maryland 1 km global land cover products. Int. J. Remote Sens. 6 (6 and 7), 1365–1373.

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A., Schmullius, C., 2008. Some challenges in global land cover mapping: an assessment of agreement and accuracy in existing 1 km datasets. Remote Sens. Environ. 112 (5), 2538–2556.

Homer, C., Huang, C., Yang, L., Wylie, B., Coan, M., 2004. Development of a 2001 national landcover database for the United States. Photogramm. Eng. Remote Sens. 70 (7), 829–840.

Johnson, D.M., Mueller, R., 2010. The 2009 cropland data layer. Photogramm. Eng. Remote Sens. 76 (11), 1201–1205.

Jung, M., Henkel, K., Herold, M., Churkina, G., 2006. Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sens. Environ. 101 (4), 534–553.

King, R.S., Baker, M.E., Whigham, D.F., Weller, D.E., Jordan, T.E., Kazyak, P.F., Hurd, M.K., 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecol. Appl. 15 (1), 137–153.

Landis, D.A., Gardiner, M.M., van der Werf, W., Swinton, S.M., 2008. Increasing corn for biofuel production reduces biocontrol services in agricultural landscapes. Proc. Nat. Acad. Sci. 105 (51), 20552–20557.

Loveland, T.R., Reed, B., Brown, J., Ohlen, D., Zhu, Z., Yand, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sens. 21 (6 and 7), 1303–1330.

Mahmood, R., Pielke Sr., R.A., Hubbard, G.K., Niyogi, D., Bonan, G., McNider, R., Lawrence, P., McAlpine, C., Etter, A., Gameda, S., Qian, B., Carleton, A., Beltran-Prezekurat, A., Chase, T., Quintanar, A.I., Adegoke, J.O., Vezhapparambu, S., Conner, G., Asefi, S., Sertel, E., Legates, D.R., Wu, Y., Hale, R., Frauenfeld, O.W., Watts, A., Shepherd, M., Mitra, C., Anantharaj, V.G., Fall, S., Lund, R., Trevino, A., Blanken, P., Du, J., Chang, H., Leeper, R., Nair, U.S., Dobler, S., Deo, R., Syktus, J., 2010. Impacts of land use/land cover change on climate and future research priorities. Bull. Am. Meteorol. Soc. 91 (1), 37–46.

Margosian, M.L., Garrett, K.A., Hutchinson, J.M.S., With, K.A., 2009. Connectivity of the American agricultural landscape: assessing the national risk of crop pest and disease spread. Bioscience 59, 141–151.

Maxwell, S.K., Wood, E.C., Janus, A., 2008. Comparison of the USGS 2001 NLCD to the 2002 USDA census of agriculture for the Upper Midwest United States. Agric. Ecosyst. Environ. 127 (1–2), 141–145.

McDonald, R.I., Fargione, J., Kiesecker, J., Miller, W.M., Powell, J., 2009. Energy sprawl or energy efficiency: climate policy impacts on natural habitat for the United States of America. PLoS ONE 4 (8), e6802.

Meehan, T.D., Werling, B.P., Landis, D.A., Gratton, C., 2011. Agricultural landscape simplification and insecticide use in the Midwestern United States. Proc. Nat. Acad. Sci. 108 (28), 11500–11505.

Meehan, T.D., Hurlbert, A.H., Gratton, C., 2010. Bird communities in future bioenergy landscapes of the Upper Midwest. Proc. Nat. Acad. Sci. 107 (43), 18533–18538.

Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem. Cycles 22 (BG1022), 19.

Nelson, G.C., Robertson, R.D., 2007. Comparing the GLC2000 and GeoCover LC land cover datasets for use in economic modelling of land use. Int. J. Remote Sens. 28 (19), 4243–4262.

Pielke Sr., R.A., Pitman, A., Niyogi, D., Mahmood, R., McAlpine, C., Hossain, F., Goldewijk, K.K., Nair, U., Betts, R., Fall, S., Reichstein, R., Kabat, P., de Noblet, N., 2011. Land use/land cover changes and climate: modeling analysis and observational evidence. WIREs Climate Change 2011 (2), 828–850.

Pielke Sr., R.A., Adegoke, J.O., Chase, T.N., Marshall, C.H., Matsui, T., Niyogi, D., 2007. A new paradigm for assessing the role of agriculture in the climate system and in climate change. Agric. For. Meteorol. 142, 234–254.

Raddatz, R.L., 2003. Agriculture and tornadoes on the Canadian prairies: potential impact of increasing atmospheric CO^2 on summer severe weather. Nat. Hazards 290, 113–122.

Rudel, T.K., Schneider, L., Uriarte, M., Turner, B.L.I.I., DeFries, R., Lawrence, D., Geoghegan, J., Hecht, S., Ickowitz, A., Lambin, E.F., Birkenholtz, T., Baptista, S., Grau, R., 2009. Agricultural intensification and changes in cultivated areas, 1970–2005. Proc. Nat. Acad. Sci. 106 (49), 20675–20680.

Schlenker, W., Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to US crop yields under climate change. Proc. Nat. Acad. Sci. 106 (37), 15594–15598.

Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., Yu, T., 2008. Use of US croplands for biofuels increases greenhouse gases through emission from land-use change. Science 319, 1238–1240.

Shao, Y., Lunetta, R.S., Ediriwickrema, J., Iiames, J., 2010. Mapping cropland and major crop types across the Great Lakes Basin using MODIS-NDVI data. Photogramm. Eng. Remote Sens. 76 (1), 73–84.

Steyaert, L.T., Knox, R.G., 2008. Reconstructed historical land cover and biophysical parameters for studies of land-atmosphere interactions within the eastern United States. J. Geophys. Res. 113 (D02101), 27.

Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I., Thies, C., 2005. Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. Ecol. Lett. 8, 857–874.

US Department of Agriculture, National Agricultural Statistics Service, 2009a. 2007 Census of Agriculture, United States Summary and State Data, volume 1, Geographic Area Series, Part 51, AC-07-A-51.

US Department of Agriculture, National Agricultural Statistics Service, 2009b. Census of Agriculture Methodology, Appendix A. <http://www.agcensus.usda.gov/Publications/2007/Full_Report/Volume_1,_Chapter_1_US/usappxa.pdf> (accessed 13.08.12).

US Department of Agriculture, National Agricultural Statistics Service, 2011. Acreage. <http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1000> (accessed 13.08.12).

US Deparment of Agriculture, National Agricultural Statistics Service, 2012. Cropland Data Layer. <http://nassgeodata.gmu.edu/CropScape/> (accessed 13.08.12).

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, J.N., 2001. Completion of the 1990s national land cover data set for the conterminous United States. Photogramm. Eng. Remote Sens. 67 (6), 650–662.

Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250m NDVI data: an assessment for the US Central Great Plains. Remote Sens. Environ. 112 (3), 1096–1116.

Werling, B.P., Meehan, T.D., Robertson, B.A., Gratton, C., Landis, D.A., 2011. Biocontrol potential varies with changes in biofuel-crop plant communities and landscape perenniality. Global Change Biol. Bioenergy 3, 347–359.

West, T.O., Brandt, G.S., Baskaran, L.M., Hellwinckel, C.M., Mueller, R., Bernacchi, C.J., Bandaru, V., Yang, B., Wilson, B.S., Maryland, G., Nelson, R., De La Terre Ugarte, D.G., Post, W.M., 2010. Cropland carbon fluxes in the United States: Increasing geospatial resolution of inventory-based carbon accounting. Ecol. Appl. 20 (4), 1074–1086.

Xian, G., Homer, C., Fry, J., 2009. Updating the 2001 national land cover database land cover classification to 2006 by using Landsat imagery change detection methods. Remote Sens. Environ. 113 (6), 1133–1147.

You, L., Wood, S., Wood-Sichra, U., 2009. Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. Agric. Syst. 99, 126–140.