

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

USGS Staff -- Published Research

US Geological Survey

2008

Comparison of the USGS 2001 NLCD to the 2002 USDA Census of Agriculture for the Upper Midwest United States

S.K. Maxwell

Arctic Slope Regional Corporation, maxwell@usgs.gov

E.C. Wood

Arctic Slope Regional Corporation, woodec@usgs.gov

A. Janus

Arctic Slope Regional Corporation

Follow this and additional works at: <https://digitalcommons.unl.edu/usgsstaffpub>

Maxwell, S.K.; Wood, E.C.; and Janus, A., "Comparison of the USGS 2001 NLCD to the 2002 USDA Census of Agriculture for the Upper Midwest United States" (2008). *USGS Staff -- Published Research*. 525.
<https://digitalcommons.unl.edu/usgsstaffpub/525>

This Article is brought to you for free and open access by the US Geological Survey at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in USGS Staff -- Published Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Short communication

Comparison of the USGS 2001 NLCD to the 2002 USDA Census of Agriculture for the Upper Midwest United States

S.K. Maxwell^{*}, E.C. Wood, A. Janus

Arctic Slope Regional Corporation (ASRC) Research and Technology Solutions (ARTS), contractor to the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA

ARTICLE INFO

Article history:

Received 2 January 2008
 Received in revised form 29 February 2008
 Accepted 6 March 2008
 Available online 14 May 2008

Keywords:

GIS
 National Land Cover Data (NLCD) map
 Cropland
 Agriculture
 Agricultural census

ABSTRACT

The U.S. Geological Survey (USGS) 2001 National Land Cover Database (NLCD) was compared to the U.S. Department of Agriculture (USDA) 2002 Census of Agriculture. We compared areal estimates for cropland at the state and county level for 14 States in the Upper Midwest region of the United States. Absolute differences between the NLCD and Census cropland areal estimates at the state level ranged from 1.3% (Minnesota) to 37.0% (Wisconsin). The majority of counties (74.5%) had differences of less than 100 km². 7.2% of the counties had differences of more than 200 km². Regions where the largest areal differences occurred were in southern Illinois, North Dakota, South Dakota, and Wisconsin, and generally occurred in areas with the lowest proportions of cropland (i.e., dominated by forest or grassland). Before using the 2001 NLCD for agricultural applications, such as mapping of specific crop types, users should be aware of the potential for misclassification errors, especially where the proportion of cropland to other land cover types is fairly low.

Published by Elsevier B.V.

1. Introduction

The U.S. Geological Survey (USGS) National Land Cover Database (NLCD) is widely used for environmental mapping and modeling (Homer et al., 2007). The land cover types classified in the NLCD are general classes such as deciduous forest, grassland, and cultivated cropland. However, more specific land cover classes such as crop type may be needed in some applications. Typically, to identify specific crop types, satellite imagery for the year of interest needs to be obtained and classified. The process of classifying specific crop types, such as corn and soybeans, from satellite imagery can be greatly simplified by first masking the imagery using a general land cover product like the 2001 NLCD to eliminate nonagricultural land cover types (Morse et al., 2004; Kastens et al., 2005; Maxwell et al., 2006).

We are currently studying crop rotation patterns in the Upper Midwest region of the United States and plan to use the 2001 NLCD to simplify the process of mapping crop types. The 2001 NLCD, which represents the most recent update of NLCD, was released in January 2007. A formal assessment of its accuracy is currently being done (Homer et al., 2007). The objective of this study was to compare the 2001 NLCD agricultural land area (i.e., “cultivated

cropland” class) to the USDA Census data in the Midwestern Corn Belt region. This paper presents the results of our evaluation.

2. Approach

2.1. Study area description

The study area extends across the entire Upper Midwest region of the United States—also referred to as the Midwestern Corn Belt region. The area spans 14 states (1229 counties) covering 2.15 M square kilometers (Fig. 1). The landscape is primarily agricultural land with heavily forested regions in the northern, eastern, and southern areas. Grassland dominates large areas of the western part of the study area in western South Dakota and Nebraska. Corn and soybeans are the dominant crops grown in the central part of the study area with areas of sunflowers and spring wheat in the northwest, winter wheat in the west and southwest, and sorghum in the southwest (NASS 2002).

2.2. The NLCD

The 2001 NLCD was developed to provide a consistent land cover product for the entire U.S. (<http://landcover.usgs.gov>; Homer et al., 2007). The product was derived from 30 m resolution multitemporal Landsat 5 and 7 imagery. Three image dates were collected for each Landsat path-row footprint to capture critical

^{*} Corresponding author.

E-mail address: maxwell@usgs.gov (S.K. Maxwell).

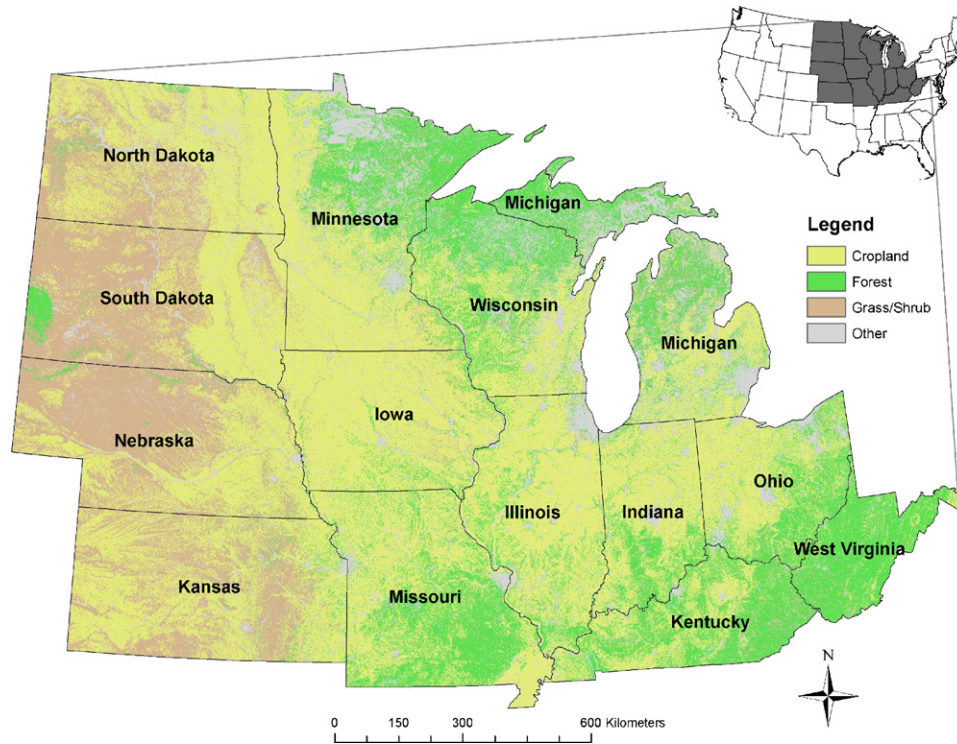


Fig. 1. Study area. Map was derived from the USGS 2001 NLCD.

stages of vegetation growth over the season (early, peak, and late vegetation green-up) (Yang et al., 2001). Images were optimally selected from the year 2001 however, when cloud-free images were not available, images were selected from the next closest year. The U.S. was stratified into 66 regions, referred to as “mapping zones”, to facilitate and optimize the classification process (Homer and Gallant, 2001). Mapping zones represent relative homogeneity with respect to biophysical and spectral characteristics. Classification was performed on each individual mapping zone using a supervised classification approach applied to a combination of multitemporal image mosaics and several ancillary data sets (e.g., terrain, population density, roads) (Homer et al., 2007). Sixteen general land cover classes were mapped for the Upper Midwest region. The specific class we evaluated is referred to as “cultivated cropland” and is described as (Homer et al., 2004):

“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.”

2.3. USDA Census of Agriculture

The Census of Agriculture is performed every five years by the USDA National Agricultural Statistics Service (NASS). The Census is conducted by sending questionnaires to every farm that had more than \$1000 of agricultural products either produced or sold, or normally would have been sold during the Census year. Approximately 2.8 M questionnaires were mailed for the 2002 Census with an 88.0% response rate (2002 Census of Agriculture). This information is summarized at the county level and published as the Census of Agriculture. Potential sources of error, such as non-response or incomplete mailing lists, are compensated for by applying an adjustment factor at the state level. These adjustments

are subsequently allocated to the county level using a weighting factor. The state level estimates are considered to be more accurate than the county level estimates (www.agcensus.usda.gov/Help/FAQs/2002_Census/index2.asp; accessed October 2007).

A section within of the survey questionnaire requests information on land use during the year of the survey in four categories: cropland, woodland, permanent pasture and rangeland, and “other land”. We used cropland acreage, also referred to as “total cropland” in the Census, in our comparison. Cropland is described as (2002 Census of Agriculture):

“Cropland harvested, cropland used only for pasture or grazing, cropland idle or used for cover crops or soil improvement but not harvested and not pastured or grazed, cropland on which all crops failed or were abandoned, and cropland in cultivated summer fallow.”

The Census cropland acreage value also includes land that is not cultivated; therefore, an adjustment was necessary to compensate for these values. The USDA “cropland” class includes acreage for pasture and hay, but in the NLCD map, pasture and hay are a separate class. We compensated for this discrepancy by subtracting the Census acreage values for hay (or forage) and pasture from the cropland value. Cropland acreage was retrieved from Census Table 1, forage acreage was retrieved from Census Table 23, and pasture acreage was retrieved from Census Table 8 (http://www.nass.usda.gov/Census/Create_Census_US_CNTY.jsp; accessed April 2007). Nursery and greenhouse crops were also included as part of the USDA cropland value; however, we did not compensate for these land areas due to the small amount of acreage they represent.

2.4. Comparison method

Comparison between the Census of Agriculture and the NLCD was performed using ArcGIS 9.2 geospatial analysis software (ESRI, 2006). NLCD cropland area for each county was calculated by

overlaying a county boundary layer on the NLCD digital map, summing the number of pixels classified as cultivated cropland, and converting to hectares (number of pixels times 0.09). Census cultivated cropland area for each county was calculated by subtracting forage and pasture acreages from total cropland acreage, and converting to hectares (number of acres times 0.4047). The Federal Information Processing Standard (FIPS) county code was used to link the county level Census data to

the geospatial map. Difference between the Census and NLCD was calculated as Census minus NLCD.

3. Results

Total cropland estimated by the Census for the entire study area was 79.0 M hectares, whereas cropland estimated by the 2001 NLCD was 77.6 M hectares (1.8% lower). The majority of counties

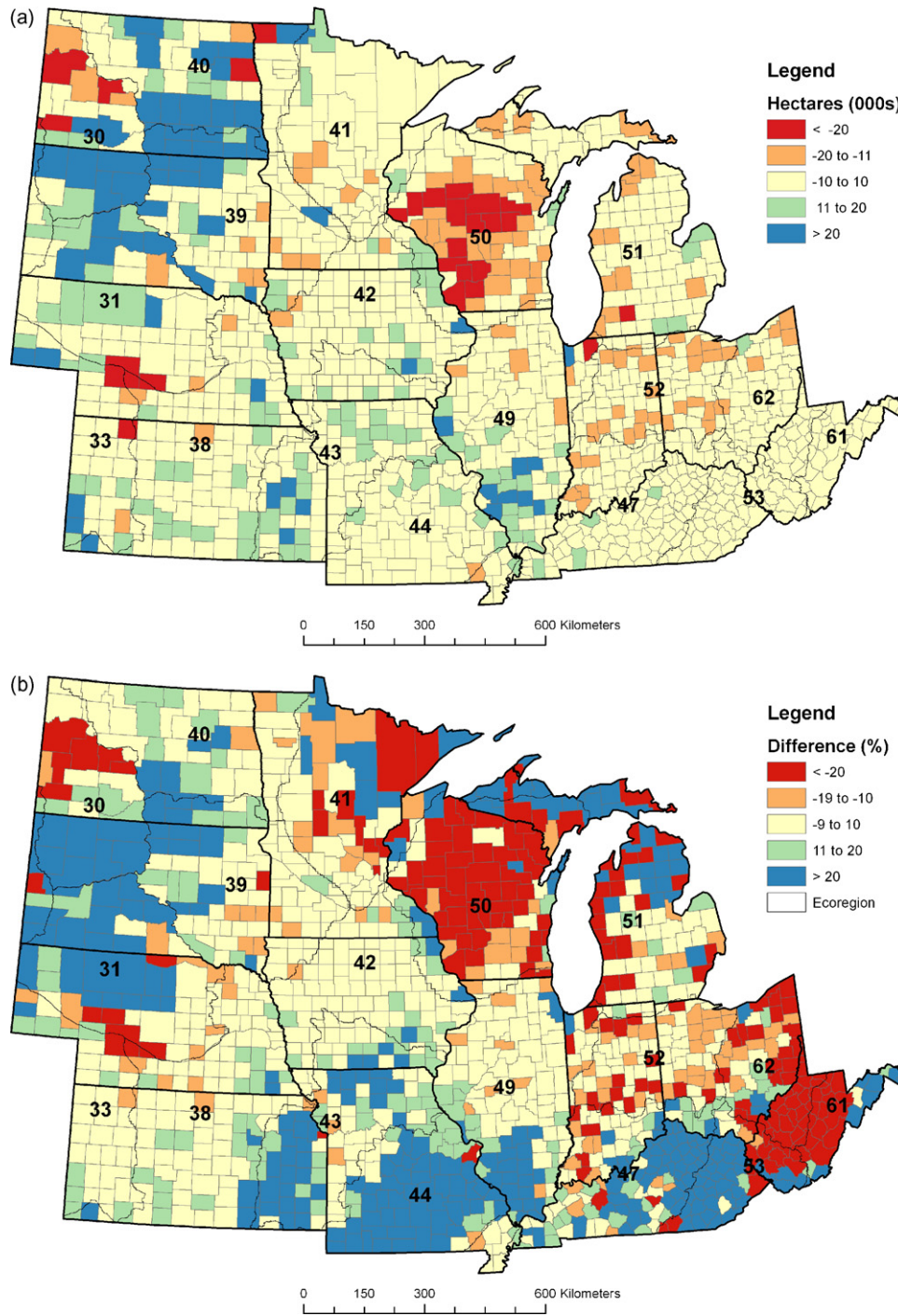


Fig. 2. (a) Hectares difference between 2002 USDA Census of Agriculture and 2001 NLCD cropland estimate at the county level. Counties represented in orange and red colors are where the NLCD estimates were higher than the Census estimate. Counties represented in green and blue colors are where the NLCD estimates were lower than the Census estimate. (b) Percent difference between 2002 USDA Census of Agriculture and 2001 NLCD cropland estimate at the county level. Counties represented in orange and red colors are where the NLCD estimates were higher than the Census estimate. Counties represented in green and blue colors are where the NLCD estimates were lower than the Census estimate.

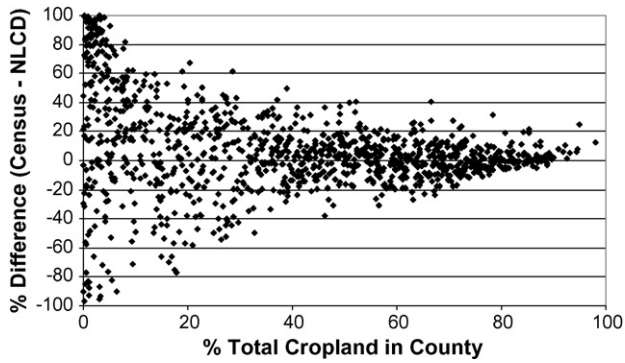


Fig. 3. Relationship between percent of county in cropland and percent difference between 2002 Census and 2001 NLCD for cultivated cropland.

Table 1

Comparison of USDA 2002 Census cropland area estimate to USGS 2001 NLCD cropland area estimate for States in the Upper Midwest U.S. (ha)

State	Census	NLCD	Difference	%Difference
Illinois	9,272,693	8,771,222	501,471	5.4
Indiana	4,789,259	5,183,478	-394,219	-8.2
Iowa	9,819,739	9,527,853	291,886	3.0
Kansas	9,848,116	9,168,601	679,515	6.9
Kentucky	1,405,490	1,158,704	246,785	17.6
Michigan	2,574,479	2,848,367	-273,888	-10.6
Minnesota	8,048,401	8,154,152	-105,751	-1.3
Missouri	4,274,946	3,823,645	451,301	10.6
Nebraska	7,297,139	7,118,967	178,172	2.4
North Dakota	9,110,347	8,533,272	577,075	6.3
Ohio	3,826,552	4,188,814	-362,261	-9.5
South Dakota	5,903,849	5,180,136	723,714	12.3
West Virginia	64,033	81,368	-17,336	-27.1
Wisconsin	2,791,169	3,824,250	-1,033,080	-37.0
Total	79,026,211	77,562,828	1,463,384	1.9

(74.5%) were less than 10,000 ha different than the Census estimate (Fig. 2a) with 7.2% of the counties having differences of more than 20,000 ha. Regional clusters of large differences in total hectares were found in Wisconsin (NLCD estimate was greater), North and South Dakota (NLCD estimates were lower), and southern Illinois (NLCD estimate was greater) (Fig. 2a).

The highest percent differences occurred along the outer edges of the study area where the lowest proportion of cropland was located (Fig. 3 (scatterplot), and compare Fig. 1 to Fig. 2b). Census estimates of cropland area in regions dominated by grassland tended to be less than estimated by NLCD. Census estimates of cropland in forest dominated landscapes, however, were less than NLCD estimates in the southern part of the study area and greater than NLCD estimates in the northern and eastern regions in the NLCD.

4. Discussion

We accepted the Census as the most accurate information available on cropland acreages; however, errors are known to exist in this dataset resulting from various issues, such as incompleteness of the mailing list and non-response of participants. Because these adjustments are applied at the state level, and subsequently apportioned to the county level where appropriate, the state level estimates are considered the most accurate (www.agcensus.usda.gov/Help/FAQs/2002_Census/index2.asp; accessed October 2007). Aggregating our data to the state level still resulted in

large differences (absolute) between the NLCD and Census estimates ranging from 1.3% to 37.0% (mean = 11.3%) (Table 1).

Discrepancies between the NLCD and Census estimates may also vary from one mapping zone to another due to differences in the individual analysts' classification approach. We found no relationship between mapping zone and the NLCD/Census and difference (Fig. 2a and b). We also evaluated the cross-validation product accuracy estimates included with the metadata. NLCD producers suggest that the cross-validation estimate can be used as an indication of map accuracy until the formal accuracy assessment has been completed (Homer et al., 2007). We found that the cross-validation values were not consistent with our results. For example, the NLCD accuracy for Zone 30 (western South and North Dakota) was stated as 69.9% in the metadata file. This estimate was consistent with our results as we did find high differences in this region between the Census and NLCD cropland estimates (Fig. 2a and b). In contrast, the overall NLCD Zone 50 (Wisconsin) accuracy is reported at 91.3%, which is not consistent with our results. We found the difference between the NLCD cropland estimate and Census estimate in this zone to be high. The NLCD cross-validation accuracy estimates were reported as overall estimates for all classes and may not be indicative of the specific cultivated crop class.

Distinguishing croplands from other land cover types such as grassland or hay/pasture can be difficult depending on the image dates selected and condition of the vegetation at the time of acquisition. In addition, although 2001 was the nominal year for acquisition of the three dates of imagery collected over the growing season, the dates may not necessarily occur in the same year. For example, a particular region may have been classified using an early season image from 2001, a peak greenness image from 1999, and a late season image from 2003. This may result in confusion between cultivated crops and other cover types in regions where crop rotation occurs (e.g., hay to cropland). Separation of these cover types can be particularly challenging for dryland crops grown in regions with low precipitation such as in western South Dakota (average precipitation <45 cm). The NLCD estimate for cropland was over 50% lower (>30,000 ha) than the Census estimate for most western South Dakota counties. Most crops in this region are nonirrigated and the greenness, or health, of the crop canopy can be severely reduced in years of drought. We suspect that misclassification errors were high between the cropland, hay/pasture, and grassland classes in this region because of the lower canopy greenness of the region's dryland crops. We recommend merging these classes prior to attempting to classify individual crop types in this region and other regions where the NLCD estimate is much lower than the Census estimate (e.g., southeastern North Dakota).

The NLCD cropland estimate was much higher than the Census estimate in several regions, such as in many Wisconsin counties (greater than 20.0% higher). The landscape in much of Wisconsin is a mosaic of forest and cropland (Fig. 1). The average farm size in Wisconsin is much smaller (81.3 ha) compared to the surrounding states of Minnesota, Iowa, and Illinois (145.1 ha average) (2002 Census of Agriculture). This region is likely to have higher classification errors in cropland because of the increased number of pixels containing a mixture of land cover types (e.g., forest and cropland).

The application of a postclassification smoothing algorithm may also have contributed to under/overestimates in cropland. A "smart eliminate" algorithm was applied to the NLCD map to aggregate the map to a 0.4-ha or larger minimum mapping unit (MMU) (Homer et al., 2007). Higher MMU thresholds were applied to cropland categories to reduce commission errors. This aggregation would eliminate miscellaneous pixels (e.g., grass

borders and drainage) within and near large contiguous crop fields resulting in overestimates of cropland. In regions where very small crop fields are intermixed within other land cover types (e.g., forest, grassland), the opposite may occur resulting in a reduction of cropland in the NLCD map.

5. Conclusion

The 2001 NLCD may provide a good baseline data layer to use for stratifying nonagricultural lands in satellite imagery prior to classifying specific crop types. We compared the NLCD map area estimates for cropland to the USDA 2002 Census estimates at the county level for 14 states in the Upper Midwest region of the United States. We found that:

- The majority of counties (74.5%) had differences of less than 10,000 ha.
- 7.2% of the counties had differences of more than 20,000 ha. The counties with the largest differences generally clustered in regions with the lowest proportions of cropland (i.e., dominated by forest or grassland).
- Differences between the NLCD and Census estimates were not associated with mapping zone (i.e., there did not appear to be an image analyst bias).

Users who apply the 2001 NLCD as a general mask for further mapping of specific crop type need to be aware of the potential for misclassification errors in the NLCD, especially where the proportion of cropland to other land cover types is fairly low. We recommend combining the cultivated land NLCD class with other classes, such as grassland and hay/pasture, in regions where

the NLCD cropland estimate is lower than the Census estimate (e.g., western South Dakota).

Acknowledgement

Work was performed under USGS contract 03CRCN0001.

References

- 2002 Census of Agriculture, Issued 2004. U.S. Department of Agriculture, United States Summary and State Data, vol. 1, Geographic Area Series Part 51, AC-02-A-51, National Agricultural Statistics Service.
- ESRI, 2006. ArcGIS 9.2 ESRI, Inc., Redlands, CA, USA.
- Homer, C.G., Gallant, A., 2001. Partitioning the conterminous United States into mapping zones for Landsat TM land cover mapping, USGS Draft White Paper. Available at: <http://landcover.usgs.gov/pdf/homer.pdf>.
- Homer, C., Huang, C., Yang, L., Wylie, B., Coan, M., 2004. Development of a 2001 National Land-Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing* 70, 829–840.
- Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., Wickham, J., 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing* 73, 337–341.
- Kastens, J.H., Kastens, T.L., Kastens, D.L.A., Price, K., Martinko, E.A., Lee, R., 2005. Image masking for crop yield forecasting using AVHRR NDVI time series imagery. *Remote Sensing of Environment* 99 (3), 341–356.
- Maxwell, S.K., Nuckols, J.R., Ward, M.H., 2006. A method for mapping corn using the US Geological Survey 1992 National Land Cover Dataset. *Computers and Electronics in Agriculture* 51, 54–65.
- Morse, A., Kramber, W.J., Allen, R.G., Tasumi, M., 2004. Use of the METRIC evapotranspiration model to compute water use by irrigated agriculture in Idaho. In: *Proceedings of the 2004 International Geoscience and Remote Sensing Symposium*, Anchorage, AK, vol. 3, pp. 2134–2137.
- Yang, L., Homer, C.G., Hegge, K., Huang, C., Wylie, B., Reed, B., 2001. A Landsat 7 scene selection strategy for a National Land Cover Database, *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium*, Sydney, Australia. Available at: http://landcover.usgs.gov/pdf/l7_scene_select.pdf.