Developing a dataset to assess ecosystem services in the Midwest, United States

Megan Mehaffey[1]*, Rick Van Remortel[2], Elizabeth Smith[3], and Randy Bruins[4]

1. Environmental Sciences Division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, 109 T.W. Alexander Dr. RTP, N.C., USA mehaffey.megan@epa.gov.

2. Lockheed Martin Information Technology RSSS, 1050 E. Flamingo Rd, Suite N-240, Las Vegas, NV, 89119, USA rvanremo@lmepo.com

3. Environmental Sciences Division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, 109 T.W. Alexander Dr. RTP, N.C., USA, smith.betsy@epa.gov

4. Ecological Exposure Research Division, National Exposure Research Laboratory, U.S, Environmental Protection Agency, 26 W. Martin Luther King Dr.,Cincinnati, OH 45268, bruins.randy@epa.gov

*To whom the correspondence may be addressed:

Megan Mehaffey
USEPA MCE243-05
109 T.W. Alexander Dr.
RTP, NC, USA
(v) 919-541-4205
(f) 919-541-4329
Email:mehaffey.megan@epa.gov

Abstract

The Midwest United States produces around one quarter of the world's grain supply. The demand for corn ethanol is likely to cause a shift toward greater corn planting. To be prepared for the potential impacts from increased corn production we need a better understanding of the current state of ecosystem services in this region. In this paper we describe a unique procedure for developing a dataset containing multiple variables useful in modeling ecological responses and tradeoffs. We demonstrate how to construct a detailed land cover classification and link it to yield and agriculture practices. We used the 2001 National Land Cover Dataset (NLCD) to spatially constrain the datasets during overlay analysis. With this method we found that the percent agreement between classifications was frequently greater than 80%, indicating little change to the original base layer accuracies. Using three different land cover datasets we were able to add 18 classes for agriculture and 155 classes of natural cover. We then added additional variables of yield, fertilizer and pesticide application rates, field residue, irrigation percentages, and tillage practices that were linked tied to the new more detailed land cover. The final Midwest dataset contained 15.5 million grid values and 15 variables. Capturing the land cover and land management information at the 30-m grid scale allows for aggregation and modeling of the ecosystem services at a variety of spatial scales. As a final step we demonstrate a tradeoff evaluation between corn yield and nitrogen loadings using our dataset. The effort required to develop the Midwest dataset was greater than initially anticipated. However, the benefit of being able to calculate derivative variables and add new variable justifies the time expenditure needed to create such a detailed database.

1. Introduction

There is an urgent need to mature our understanding of the services provided by the ecosystems of the Midwestern United States. Ecosystem services have been variously defined as the benefits people obtain from ecosystems (Millennium Ecosystem Assessment, 2005) and aspects of ecosystems utilized to produce human well-being (Fisher et al., 2009), and they include, for example, the provision of clean air, clean water, flood control and nature-based recreation opportunities, as well as the production of food, fuel and fiber. The Midwest is responsible for a significant proportion of the world's grain production. For example, in 2005, approximately 23 percent of the world's maize, soybeans and wheat originated from the 12-state area outlined in Figure 1. However, the Midwest also exports less desirable products in tandem with grain. Nutrient runoff from Midwest farmlands contributes to seasonal hypoxia in the Gulf of Mexico (Alexander et al., 2008) and eutrophication of local streams and lakes, while commonly used herbicides are frequently detected in shallow groundwater (Gilliom and Hamilton, 2006). Farming within the rich floodplains of the Midwest has modified the drainage and water-holding capacity of these soils resulting in increased heights and frequencies of floods in the Upper Mississippi River basin (Pinter et al. 2006).

Water quality and quantity problems are expected to be exacerbated as rising grain prices spur a switch to monoculture cropping of corn into lands with competing crops (i.e. cotton and wheat), and potential conversion of Conservation Reserve Program (CRP) lands back to field crops (Westcott, 2007). The pressure to increase corn production is also likely to affect already marginalized wildlife species reducing populations, thereby decreasing important ecosystem services such as the existence of native biodiversity, wildlife viewing opportunities and recreational use.

Understanding the tradeoffs in services expected to occur in the Midwest as a result of the increased demand for ethanol could help determine a better way of managing future resources for maximized benefits. Secretary of the United States Department of Agriculture (USDA) Mike Johanns stated in 2005, "I see a future where credits for clean water, greenhouse gases, or wetlands can be traded as easily as corn or soybeans (USDA, 2005). However, achieving this goal will require the collection and compilation of enough information to create meaningful models and maps for these ecosystem service tradeoffs".

Currently available map products are targeted to meet interest group or agency specific needs. For example, the USDA National Agricultural Statistics Service (NASS) tracks production and management of crops at the state and county level. The USDA has also created the annual Cropland Data Layer (CDL) land cover with an eye to monitoring changes in acreage for major crop types (USDA CDL, 2008). Federal agencies including the United States Geologic Survey (USGS) and U.S. Environmental Protection Agency (EPA) generated the 2001 National Land Cover Database (NLCD) products, which is a medium resolution geodatasets intended to support national and regional assessments of land cover and use change (Yang, 2008). Other spatial mapping efforts are strongly supported by non-profit groups such as The Nature Conservancy and NatureServe; for instance, the LANDFIRE program was developed to map vegetation fuel loads for modeling the spread of fire through wildlands (LANDFIRE, 2007), while the Gap Analysis Program (GAP) has been focused on habitat for modeling species populations (Jennings, 2000).

To begin the process of expanding our understanding of tradeoffs in ecosystem services we needed to construct a spatial dataset for the Midwest. The dataset needed to contain both a detailed classification of agricultural practices and more specific classifications for natural cover. The additional information would be used to model or calculating ecosystem functions and serves as the cornerstone for creation of an alternative futures landcape. The future landscape would represent change resulting from the increased corn production expected to satisfy future 2022 ethanol fuel mandates. In this paper, our process for integrating data from multiple sources into a new more detailed dataset is set described (Figure 2).

2. Methods

We delineated a boundary based on the USGS 8-digit hydrologic units (HUCs) intersecting the 12 most productive Midwest states in the U.S (Figure 1). This study area provided a contiguous landscape encompassing large portions of the Midwestern plains and prairie ecoregions (Omernik 1987). The ecoregions are home to at least 34 threatened and endangered species The 12 states also include 165 out of 211 operational US bioethanol production facilities as of July 14, 2009 making them the primary location for increased corn planting to meet the demand for ethanol production (Renewable Fuels Association, 2009). Data and developed datasets were compiled for this area.

2.1 Data

National land cover database (NLCD)

We used the 2001 NLCD database as our base layer in preparing our augmented land cover dataset for the Midwest. The 2001 NLCD used an improved classification algorithm to develop 16 land cover classes for the conterminous United States (Homer et al., 2004). The NLCD is a 30-meter product that was derived from satellite imagery from Landsat 5 and 7. The NLCD provides a consistent coverage that can be used on regional to national scale. The dataset was developed with the idea that many users would want to develop value added products for specific applications. We downloaded the data corresponding to our study area from the MRLC multizone site (http://www.mrlc.gov\nlcd_multizone_map.php).

LANDFIRE existing vegetation database

The NLCD-2001 contains only eight natural vegetation land cover types; three for forest, one for natural grassland, and two each for wetland and water. To expand the number of natural cover classes we relied on the existing vegetation type (EVT) dataset produced by the interagency LANDFIRE program. Relative to NLCD data, the 30-m-resolution EVT grid product offers a much wider variety of vegetative cover classes. The EVT was created primarily from interpretations of 2000-2002 Landsat satellite imagery and an extensive field reference dataset (LANDFIRE, 2007). In the Midwest there was a total 155 different LANDFIRE classes, which included multiple grassland/prairie, deciduous and evergreen forest, scrub/shrub and wetlands types.

Cropland data layer database

We chose the NASS Cropland data layer (CDL) dataset to expand the number of NLCD's 'cultivated crops' classes determine crop rotation. For the Midwest there were up to 56 potential classes of crops including major grains, cover crops, and vegetable and fruits. The CDL program utilizes spring and summer seasonal satellite imagery to produce crop-specific, categorized, and georeferenced output products and provide annual acreage estimates for major agricultural commodities (Mueller and Ozga, 2002). To create rotational practices we collected CDL data from 2004-2007. A fundamental change in the types of source satellite imagery used by NASS for the CDL crop type classification occurred during the specified 2004-07 time period. Anticipating a loss of available 30- meter satellite data the National Agriculture Statistics Service (NASS) made a shift from Landsat Thematic Mapper (TM) data to the 56-meter satellite imagery from the Advanced Wide Field Sensor (AWiFS).

SSURGO crop yields

The detailed land covers did not provide any indication of crop yield so we turned looked at other types of data for this resource. We determined that the highest resolution of crop yield data available to us was contained within the Natural Resource Conservation Service's (NRCS) Soil Survey Geographic Database (SSURGO) soil map unit database (NRCS, 1995). A total of 1,142 discrete SSURGO soil survey area databases were procured or downloaded from NRCS to encompass the vast majority of the Midwest study area (Soil Survey Staff, NRCS, http://soildatamart.nrcs.usda.gov). The ‘ccrpyd’ component crop yield table within each database was compiled from yield estimates generally calculated by NRCS agronomists knowledgeable about the specific soils and yields in the area. To the best of our knowledge, SSURGO crop yield estimates were routinely accuracy-screened by NRCS state agronomists for most if not all survey areas prior to their inclusion in the SSURGO digital database, which would link the estimates to a late-1990s or later lineage.

County and agricultural district annual reports

At the time we were developing our dataset SSURGO had not been completed for the whole of the Midwest and we included the county level NASS data to fill in for those locations without yield. For our purposes, annual report data from the 2004-07 time periods were downloaded by state and year from NASS and separated into tables by county and agricultural district entities (USDA NASS, 2004-2007). One of the ongoing responsibilities of NASS is to prepare and publish, in conjunction with each state's agriculture department, annual county crop production data to support USDA's farm and cooperator programs. The county data within these reports were based on a statistical sampling of farms and ranches, and agricultural districts are defined statistical groupings of counties within each state that are typically lumped according to geography, climate, and cropping practices.

Irrigated lands

The SSURGO and NASS yield values are developed for irrigated and non-irrigated crops. In order to determine the appropriate yield values we needed a spatial representation of irrigated lands within the Midwest. After reviewing several available datasets, we selected a 2001 irrigation map developed by the Center for Sustainability and the Global Environment at the University of Wisconsin having a moderate spatial resolution of 463 meters (Ozdogan and Gutman, 2008). The mapping product was derived from remote sensing radiometric data collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument and was augmented by extensive ancillary sources of global climatic and agricultural array data. The map represents the proportion of land within each grid cell that was irrigated. We used a 50 percent cutoff value for determining which yield value to use from the SSURGO and NASS datasets.

State level tillage operations and fertilizer/pesticide applications

We determined that Internet-downloadable Agricultural Resource Management Survey (ARMS) data addressing crop production practices would provide the needed information about the current status and trends in management practices for the large-acreage field crops of corn, wheat, and soybeans. Sponsored jointly by NASS and USDA's Economic Research Service (ERS), the annual ARMS data are the primary source of information relating to the financial condition, production practices, and resources use of American farm households. The ARMS program annually collects a representative field-level sampling of information on chemical applications, tillage systems, and pest, nutrient, and crop residue management practices.

2.2 Crop rotation classes

In Gassman et al.'s 2006 Upper Midwest basin mapping effort, field crops including corn, soybean, wheat, and alfalfa in conjunction with Conservation Reserve Program (CRP) and pasture lands were digitally linked to the NLCD 2001 data using simple GIS-based raster overlay processes. For our study we expanded the method described by Gassman et al. 2006 to include additional crop types and rotations.

In order to develop the crop rotation classes we consolidated 4-years of CDL data for our study area. The 4 years of CDL (2004-2007) data spanned the transition period between the uses of TM to AWiFS sateillite. We determined that retaining the 56-m resolution of the later AWiFS imagery resulted in more congruent overlays between years, and yielded fewer instances of grid cells exhibiting non-field-crop values.

We then created a 'hybrid' reclassification of the 2004-07 CDL field crop classes so that all available CDL crop cover classes could be matched across the 4-year period. To accomplish this, earlier years' classes were crosswalked and renumbered as necessary to correspond with the 2007 standard CDL class-numbering scheme. Each state's hybrid-classed field crop grids were then spatially combined across the four overlay years of 2004-07. A character attribute was added to the combined attribute table to serve as a unique identifier key indicating the actual four-year classes contributing to the final crop rotation assignments. The combination of CDL hybrid classes spanning the four overlay years yielded 96,295 distinct crop combinations within the study area.

Next, we defined a set of crop rotation classes that would adequately encompass and represent the principal crops of the region. Grid cell counts (i.e., total number of 56-m grid cells occupied for each rotation combination) were used to examine general cropping patterns. A crop category variable was enlisted to aid these assignments. Corn, soybean, wheat, alfalfa/hay, and cotton were the key crops, while miscellaneous grain and fallow/idle classes were used to distinguish among other secondary crops. A variety of combinations drawn from these crop types yielded 19 discrete crop rotation classes plus a '0' class denoting both non-field-crop and missing data situations (Table 1).

Finally, each cell in the Midwest study area was assigned a crop rotation class by iteratively looping through a 7-level subsetting hierarchy aml. Subset groups included monoculture, two crop rotations, mixed cropping, and other miscellaneous crops. The first iteration selected for monoculture occurrences. That is, where a single crop category occurred in all 4 years or where there was 3 years of single crop with no data for the other year. The second and third iteration searched for and assigned mixed rotation classes where different crop types occurred between the four years. The fourth iteration applied a 2006-07 bias for monoculture or rotational classes (i.e., rotation was assigned based on 2006-07 crops identified). The fifth and sixth iteration focused on selections and assignment of mixed fallow/idle combinations. The last iteration addressed any leftover rotational combinations which were assigned en masse to the other crop /fallow idle class.

2.3 The Midwest classification

The CDL and EVT grids values were added to the dataset using overlay analysis. The grids were linked to the NLCD by locking in their corners. New columns for each classification were then added to the NLCD attribute table. The 56-m CDL values were assigned to the 30-m grid values using center point intersection rules.

Once the classifications were joined into a single attribute table we used Statistical Analysis Software (SAS) code to construct a new classification column. We used if/then rules to assign the classes for the new Midwest classification as follows: 1) where NLCD was urban, water, and barren the values were maintained; 2) where NLCD was natural cover (i.e., forest, shrub, grass/herb, wetland) then the LANDFIRE EVT classification value was added; 3) where NLCD designated a grid cell as agriculture (i.e., class 81 or 82) we inserted replacement values from the CDL crop rotation classification.

In two cases we made exceptions were made to the classification rules. The first was where NLCD designated a grid as grassland/herbaceous (class = 71) and CDL designated the grid cell as wheat or wheat-rotation. In this case we chose to use the CDL class instead of the LANDFIRE. We based this decision on the tendency of NLCD to underestimate wheat crops in the Upper Midwest (Maxwell et al. 2008). The second exception we made was where NLCD designated a grid cell as barren but EVT identified the grid cell as recently logged. In this case we retained the EVT logged value.

After implementing the above SAS processing steps the final Midwest land cover classification contained 18 classes for agriculture and 155 classes of natural cover, 3 urban, barren and water (Figure 3). The land cover source data information was retained within the attribute file along with the new Midwest classification to insure that users would have access to the data for comparison or for development of other classifications rules.

2.4 Adding yield and management variables

The SURRGO soil map unit spatial coverage and the 'ccrpyd' component crop yield table were extracted from the soil survey file structure. Both irrigated and non-irrigated yields were area-weighted by the relative percent of area represented by each component in order to derive aggregate yield values for each soil map unit. We transformed the SSURGO soil map unit polygons into a 30-meter grid arrays and the associated crop yield data into an Arc Info attribute table. The tables for each soil map unit were subsequently joined to the appropriate grids and then all grids were merged together for the study area. We prepared SSURGO crop yield estimates for corn, soybean, wheat, and alfalfa/hay.

Where the spatial SSURGO coverage was present but there was no component crop yield table available (i.e., the states of North Dakota, Kansas, and Missouri), a series of state-specific crop productivity indexing approaches were used to derive aggregate crop yield values for the soil map units. The criteria used in those productivity indices were provided directly to us by the respective NRCS state office staffs. In other places where SSURGO data were unavailable from NRCS (about 4 percent of study area), the crop yield data were left blank (i.e., zeroed out) with the intention that potential users of the data could fill in values for these areas either from the NASS county-level data or from other specific sources. Records having null yield values (yield=0) were discarded and the remaining records were averaged across years to derive mean crop yield estimates for the two entities. Data were often reported by NASS separately for irrigated and non-irrigated conditions. However, where this wasn't the case a total crop value was given instead. Grids of 30-meter resolution were then prepared for the SSURGO and NASS yield data. The attribute tables were populated with the relevant mean annual crop yield estimates for corn, soybean, wheat, and alfalfa/hay.

Crop production practices data were downloaded and summarized by crop type and year for each state, encompassing the corn, soybean, spring wheat, and winter wheat crop types (ARMS, 2002-2005). Available ARMS tillage operations data were used to estimate tillage parameters for each crop type that included the percentage of crop residue at the time of planting, the number of tillage operations performed over the season, and the percentage of crop area cultivated for weed control. The principal weighting factor in this tillage analysis was a tillage-type variable derived from the mean percentage of five reported tillage types: no-till, ridge-till, mulch-till, reduced-till and conventional-till. The available ARMS chemical fertilizer and pesticide application data were then used to estimate annual application parameters for each crop type including nitrogen, phosphate, potash, herbicide, pesticide, and total pesticide, all expressed on a mass-per-unit-crop-area basis for each state.

After implementing the above processing steps the data were added to the expanded Midwest dataset using overlay analysis. The Midwest attribute file now contained three land cover classifications, yield data by crop type from both SSURGO and NASS, and variables for cropping practices, and pesticide use. As a last step irrigation percentages were added to the attribute table using overlay and center point rules. The final Midwest dataset contained 15.5 million grid values and 15 variables (Figure 2).

3. Results

3.1 Cropland evaluation

At the time of this paper the 2001 NLCD accuracy assessment was in press (Wickham et al., in press). The overall user accuracy for the 2001 NLCD regions (i.e. Regions IV, V, and VI) of this study ranged between 80 and 84 percent. However, the NLCD cropped agriculture class had user accuracies closer to 90 percent. The USDA CDL data publishes its accuracy assessment for each state as part of the disks/download information. The accuracy assessment of the CDL differed by state and year with the lowest agreement occurring in the year 2006 in North and South Dakota

(Table 2). USDA NASS has strived to keep accuracy of the major crops (corn, soybean, wheat) greater than 90%. In our center point overlay method we did not alter the location of the cropped agriculture from that of the NLCD. Therefore, the accuracy of cropped fields would be the same as those designated for the appropriate NLCD regions.

To determine how frequently the two land cover classifications agreed we compared agriculture classes from the combined 2004-2007 CDL crop rotation to the 2001 NLCD (Table 3). We reclassed the agriculture classes for both land covers into one combined class called field crops and created 30-m grids to run the comparison. We found that field crop classification in the CDL grid agreed with the NLCD more than 80% of the time. When we use the number of 30-meter cells for all land cover types in the study area (i.e., 2.57 x 10^9 cells) we see that the NLCD-2001 land cover grid identified lands as being agriculture somewhat less frequently than did the CDL crop rotation grid (i.e., 35.3% vs. 40.3%, respectively).

3.2 Natural cover evaluation

The LANDFIRE EVT accuracy assessment for the Midwestern states had not yet been published at the time of this study. However, as in the case of agriculture classes we controlled the spatial distribution of the natural areas using NLCD natural cover classes. Therefore, misclassification into classes of human use groups (urban, cropped, agriculture) would be expected to remain at the rate of the error found in the NLCD.

To get a measure of agreement between classes we compared the LANDFIRE natural cover classes incorporated into the Midwest classification to the NLCD. We aggregated the 155 natural cover classes into NLCD "like" groups (i.e. evergreen and deciduous trees, herbaceous and woody wetland, grasslands, and barren), then compared differences between the two grids. We summed the pixels in each natural class for the two different classifications and then calculated percent agreement (Table 4). Bareground, grasslands, woody wetland, and deciduous forest classes in the Midwest classification had the highest agreement, around 80% or greater. Herbaceous wetland agreement was slightly lower (56%) due to crossover with woody wetlands. The remaining forest class pixels were distributed across evergreen, mixed, and deciduous forest, resulting in agreement only slightly greater than 25%. The lowest agreement was in shrub/scrub with the majority of the pixel count falling in either NLCD forest or grassland classes.

4. Discussion and conclusion

With the current focus in the scientific community on ecosystem services, environmental data is needed which can be used by multiple groups for a variety of purposes. One way to meet the needs of the various modelers and data users is to combine the disparate data from the many different agencies and organizations into a single comprehensive dataset. In this paper we described a procedure for building a spatially detailed dataset containing both land cover and land use management information.

We relied on expert judgment to develop our land cover for the Midwest. In the final dataset we retained the original classifications from the CDL, LANDFIRE, and NLCD as unique variables along with the newly developed Midwest classification. By retaining the original classifications scientists have the opportunity to make comparisons between the land cover classifications or to build their own classification. The classification we developed for the Midwest was constrained spatially by the NLCD. We made no changes to water and urban classes so they would retained the Level I accuracy of the NLCD. We compared the CDL crop rotation to the NLCD and found there was 84% agreement in identification of agriculture in Midwest landscape. Using multiple years of CDL data in conjunction with the NLCD we were able to include 18 new categories for crop rotation to our classification. We chose to favor the classes of the LANDFIRE while retaining the total area of natural cover from the NLCD. As a result there was greater mixing between natural classes in the final Midwest classification. See and Fritz (2006) proposed use fuzzy logic incorporate expert judgment and evaluated disagreement between map products. However, even with only using best professional judgment we found that agreement between classifications of bareground, grassland, deciduous forest, and woody wetland remained high for the Midwest land cover.

Using various processing steps we were able to successfully add data for yield, fertilizer and pesticide application rates, field residue, irrigation percentages, and tillage practices into the final dataset. The purpose of combining land cover with management data into a 30-m grid was to provide a means to investigate ecosystem response modeling at a finer spatial resolution. Several other studies have used similar overlay, regression or dasymetric disaggregation methods to include larger scale information on agricultural practices into land cover datasets for modeling pollutant and pathogen inputs (Secchi et al. 2009; Comber et al. 2008; Cardille and Clayton 2007) With the combined Midwest land cover we can now conduct habitat suitability modeling using the mid-scale ecological unites provided by LANDFIRE (Comer et al. 2003) rather than the more general categories of the NLCD. The more detailed classification would be particularly beneficial for mapping habitat scarcity, diversity, and rarity as measures of ecosystem services.

By including crop rotation, irrigation, and application rates in the final dataset we can now estimate ecological response functions related to nutrient and pesticide loading and retention rates at a pixel scale. In addition to calculating functional response the new Midwest dataset allows for comparing tradeoffs between economic goods and ecological exposures. In the Midwest corn production is a major economic good but as demonstrated by several studies (Alexander et al. 2008, Gassman et al. 2006,) it is also strongly correlated with increase water pollution. The expected tradeoff that would occur in the Midwest is demonstrated in Figure 4. Another way to display this data would be to calculate the ratio of corn production to nitrogen load and display it in a map (i.e. by hydrologic unit) so that decisions makers could see where the highest cost in terms of nitrogen load is likely to occur in the Midwest. With very little effort estimates of farmer income for corn production and fertilizer application costs could also be added to the dataset enabling users to conduct a cost/benefit analysis.

Our final Midwest database contained over 15 million distinct objects resulting from the spatial compilation of the three land cover classifications (NLCD, CDL, LANDFIRE,), the MODIS-based irrigated lands, and SSURGO soil map unit crop yields; NASS county/district-level crop yields; and ARMS state-level tillage practices and fertilizer/pesticide applications. The effort required to develop the Midwest dataset was greater than initially anticipated. Disparate data sources, gaps in the source data, and software limitations with respect to the number of unique data layer combinations that could be handled, contributed to the time needed to compile the data. However, now that it is finished analysts are free to add to the variables presented. The benefit of the method we described in this paper is that as long as the original grid values are retained any number of derivative variables can be calculated and new data can added from other sources.

Future efforts will examine the accuracy of the small changes to the original classifications that were made for constructing the Midwest land cover.   The 2001 Midwest classification will be used to develop a future biofuels-driven 2022 scenario based on changes predicted by a linked set of econometric models which generate crop production acreages for major crops of the Midwest (i.e., corn, soybean, wheat). We also plan to add variables which will allow evaluating the changes in ecosystem response, tradeoffs and cost/benefit analysis related to the application of the herbicide Atrazine.   The procedure used in this paper for combining datasets is being applied at a national scale for use in assessment of water availability and quality, carbon stocks, and nitrogen flux.

NOTICE: *The United States Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research described here under contract number EP-D-05-088 to Lockheed Martin.*

5. References

Alexander, R.B., Smith, R.A., Schwarz, G.E., Boyer, E.W., Nolan, J.V., and Brakebill, J.W.. 2008. Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin. Environmental Science and Technology 42(3):822-830.

ARMS (2002-2005) Tillage and Fertilizer/Pesticide Homepage of Agriculture Resources Management Survey. Applications: http://www.ers.usda.gov/Data/ARMS/CropOverview.htm [Accessed : April 8, 2009].

Cardille, J. A. and Clayton, M.K. 2007. A regression tree-based method for integrating land-cover and lan-use data collected at multiple scales. Environmental and Ecological Statistics 14: 161-179

Comber, A., C. Proctor, S. Anthony. 2008. The creation of a national agricultureal land use dataset: Combining pycnophylactic interpolation with dasymetric mapping techniques. Transactions in GIS 12: 775-791.

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K. Snow, and J. Teague. 2003. Ecological Systems of the United States: A Working Classification of U.S. Terrestrial Systems. NatureServe, Arlington, VA. 75 p.

Fisher, B., Turner, R.K., Morling, P.. 2009. Defining and classifying ecosystem services for decision making. Ecological Economics 68:643-653.

Gassman, P.W., Secchi, S., and Jha, M. 2006. Upper Mississippi river basin modeling system Part1: SWAT input data requirements and issues. Coastal Hydrology and Processes (ed. By V.P. Singh and Y. J. Xu) pp 103-115. 2006 Water Resources Publications, LLC, Highlands Ranch, CO, USA.

Gilliom, R.J. and Hamilton, P.A.. 2006. Pesticides in the Nation's Streams and Ground Water, 1992–2001-- A Summary. Fact Sheet 2006-3028. U.S. Geological Survey.

Homer C., Huang C.,Yang L, Wylie, B. and Coan M. 2004. Development of a 2001 National Land-cover Database for the United States. Photogrammetric Engineering and Remote Sensing 70:829-840.

Jennings, M.D., 2000. Gap analysis: concepts, methods, and recent results: Landscape Ecology, v. 15, p. 5-20.

LANDFIRE. (2007, January - last update). Homepage of the LANDFIRE Project, U.S. Department of Agriculture, Forest Service; U.S. Department of Interior. Available at http://www.landfire.gov/index.php [Accessed : April 8, 2009].

Maxwell, S.K., Wood, E.C., and Janus, A. 2008. Comparison of the USGS 2001 NLCD to the 2002 USDA Census of Agriculture for the Upper Midwest United States. Agriculture, Ecosystems and Environment. 127:141-145.

Mueller, R., and Ozga, M., 2002. Creating a cropland data layer for an entire state. Internal research report. U.S. Department of Agriculture, National Agricultural Statistics Service, Statistical Analysis Research Center, Fairfax, VA.

Millennium Ecosystem Assessment, 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, DC.

NRCS, 1995. Soil Survey Geographic (SSURGO) Data Base: Data Use Information. Misc. Publ. 1527. U.S. Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE, USA.

Omernik, J.M., 1987, Ecoregions of the conterminous United States (map supplement): Annals of the Association of American Geographers, v. 77, no. 1, p. 118-125

Ozdogan, M, and Gutman, G., 2008. A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental U.S. Remote Sensing of Environment 112(2008):3520–3537. [MODIS-derived Irrigated Lands: http://www.sage.wisc.edu/mapsdatamodels.html, {Accessed on line March 15, 2009]

Pinter, N., Ickes, B.S., Wlosinsk, J.H. i, and van der Ploeg, R.R.. 2006. Trends in flood stages: Contrasting results from the Mississippi and Rhine River systems. Journal of Hydrology 331(3-4):554-566.

Renewable Fuels Association 2009. Biofinery Locations. Available at http://www.ethanolrfa.org/industry/locations/ [Accessed July 25, 2009]

Secchi, S., P.W. Gassman, J.R. Williams, and B. A. Babcock. 2009. Corn-based ethanol production and environmental quality: A case of Iowa and the conservation reserve program. Environmental Management 44:732-744.

See, L.M. and S. Fritz. 2006. A method to compare and improve land cover datasets: Application to the GLC-2000 and MODIS land cover products. IEEE Transactions of Geoscience and Remote Sensing. 44:1740-1746

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for [Midwest States]. Available at http://soildatamart.nrcs.usda.gov [Accessed: Februrary 15, 2009].

USDA. 2005. Remarks As Prepared by The Hon. Mike Johanns, Secretary U.S. Department of Agriculture, Innovations In Land and Resource Governance, White House Conference on Cooperative Conservation - St. Louis, MO - August 29, 2005. USDA Press Office Release No. 0335.05. U.S. Department of Agriculture.

USDA CDL (2008). Homepage of Cropland Data Layer, U.S. Department of Agriculture, National Agriculture Statistics Service, Research and Development Division, Available at: http://www.nass.usda.gov/research/Cropland/SARS1a.htm [Accessed: Feburary, 21, 2009]

USDA NASS, 2004-2007. National Agriculture Statistics Service, Crop Yields for Corn, Soybean, Wheat, Hay. Available at http://www.nass.usda.gov/Data_and_Statistics/County_Data_Files/ [Accessed April 28, 2009]

Westcott, S. 2007. USDA Issues $1.8 billion in Conservation Reserve Program rental payments. United States Department of Agriculture News Release No. 0276.07

Wickham, J.D., Stehman, S.V., Fry, Smith, J.H., and Homer, In press Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. Remote Sensing of Environment.

Yang, L., 2008, Development of the United States National Land Cover Database–experience from 1992 and 2001 implementation, in Silk road for information from imagery, Beijing, China, July 3, 2008-July 11, 2008, The International Archives of the Photogrammetry, Remote Sensing, and Spatial Information Sciences, XXXVII, pt. B4: International Society for Photogrammetry and Remote Sensing, p. 1461-1465.

Figure Titles

Figure 1. The 12 states of the Midwest study area (Illinois (IL), Indiana (IN), Iowa (IA), Kansas (KS), Michigan (MI), Minnesota (MN), Missouri (MO), Nebraska (NE), North Dakota (ND), Ohio (OH), South Dakota (SD), Wisconsin (WI)) and corresponding hydrologic unit boundary. Black circle is location of close up landcover image in Figure 3.

Figure 2. A detailed relational flowchart for development of an enhanced Midwest dataset for evaluation of ecosystem services.

Figure 3. A close up view of the differences in classifications between the original NLCD 2001 land cover and our expanded Midwest land cover. The figure shows 34 of 176 total classes in the Midwest landcover. Location is indicated by the black dot in Figure 2.

Figure 4. An example of the type of ecosystem service tradeoff assessments possible with the Midwest dataset. The scatter graph provides a way to measure the cost of increased corn yield as a function of nitrogen loading to streams. Data is sum of corn yield and nitrogen loads for twenty four thousand 12-digit hydrologic units.

Table 1. Major crops and rotational classes of the Midwest land cover

Rotation #	Rotation class	Midwest Study Area %	Field Crop Area %
0	Non-field-crop / No data	59.73	n/a
1	Corn	4.83	12.00
2	Soybean	0.10	0.24
3	Wheat	2.46	6.10
4	Cotton	0.03	0.07
5	Alfalfa_Hay	1.29	3.20
6	Fallow_Idle (includes CRP)	3.14	7.80
7	Corn / Soybean	15.15	37.61
8	Corn / Wheat	1.01	2.50
9	Corn / Other crop	0.99	2.45
10	Corn / Fallow_Idle	0.35	0.86
11	Soybean / Wheat	1.84	4.56
12	Soybean / Other crop	3.64	9.03
13	Soybean / Fallow_Idle	0.21	0.51
14	Wheat / Other crop	2.07	5.14
15	Wheat / Fallow_Idle	1.18	2.92
16	Cotton / Other crop	0.03	0.08
17	Alfalfa_Hay / Other crop	0.29	0.71
18	Miscellaneous grain / Fallow_Idle	0.89	2.20
19	Other crop / Fallow_Idle	0.80	1.99
	Totals:	100.00	100.00

Table 2. Overall accuracy assessment for cropland data layers for the 12 states in the Midwest study area from 2004 to 2007.

States	2007	2006	2005	2004
Iowa	97.2	83.0	87.3	92.5
Illinois	97.6	82.3	91.7	95.0
Indiana	95.4	79.4	92.6	82.9
Kansas	88..0	83.0		
Michigan	92.2	82.0		
Minnesota	94.8			
Missouri	91.6	87.0	89.7	94.6
North Dakota	81.1	70.0	71.1	81.9
Nebraska	92.7	94.0	74.0	82.5
Ohio	89.6	91.0		
South Dakota	83.9	61.0		
Wisconsin	90.0	89.0	77.3	89.4

Table 3. Agreement comparison between the National Land Cover Dataset and Midwest classification for field crops verses other classes.

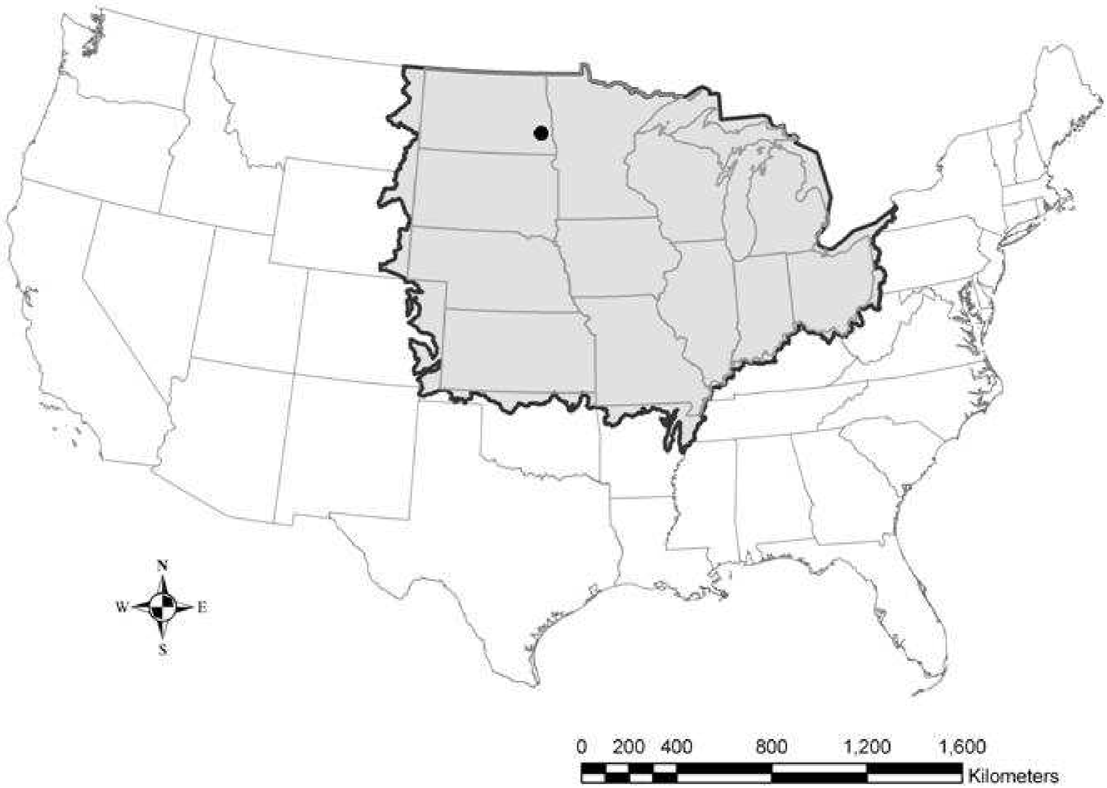
	2004-2007 Midwest Crop Rotation Classification		
2001 NLCD Classification	Field Crops	Other Class	Percent Agreement
Field Crops	**759,125 x 1000**	149,587 x 1000	84%
Other Class	277,372 x 1000	**1,387,421 x 1000**	83%
Percent Agreement	73%	90%	

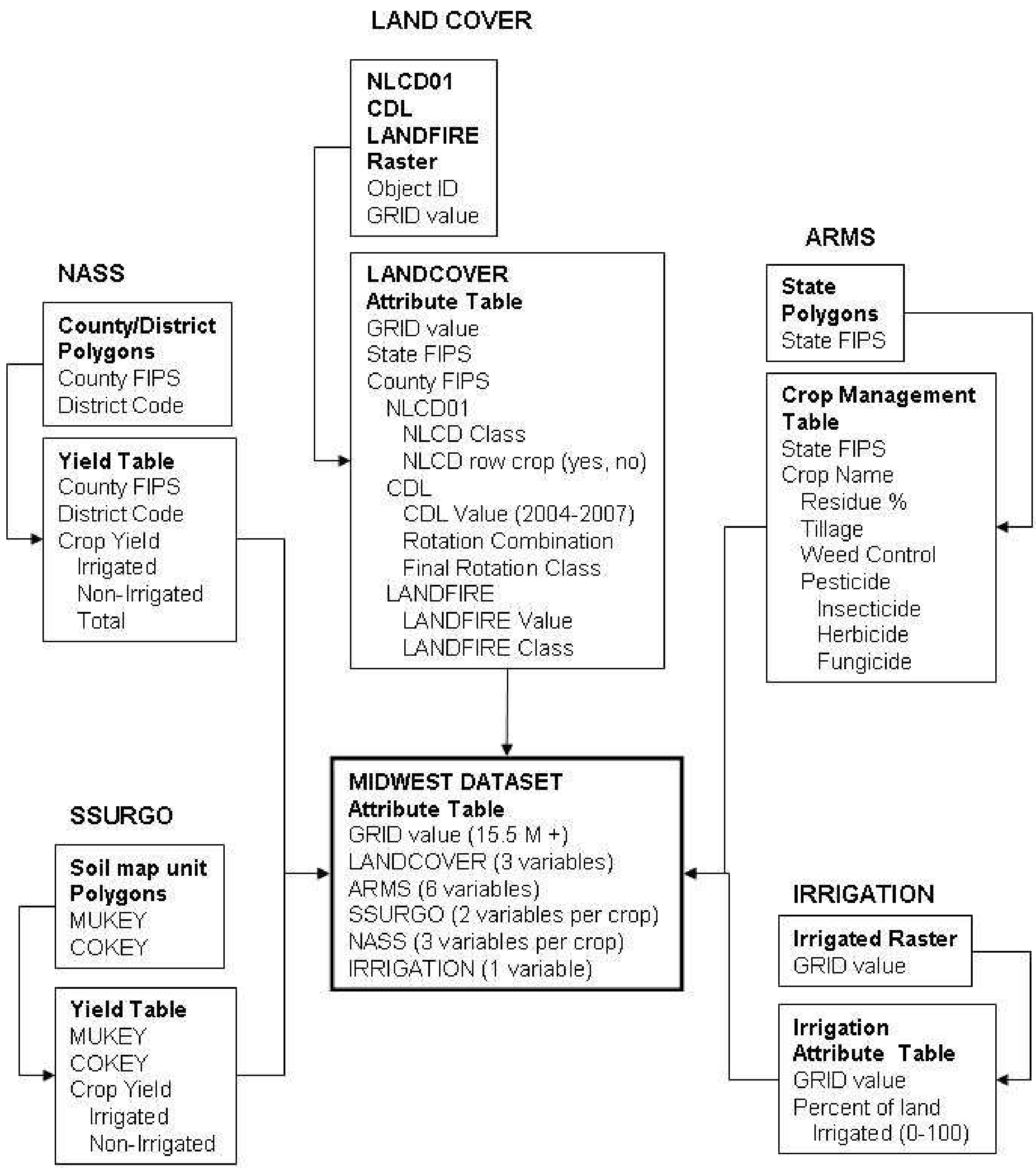
Overall Agreement (2,146,546 x 1000/2,573,505x1000) = 83%

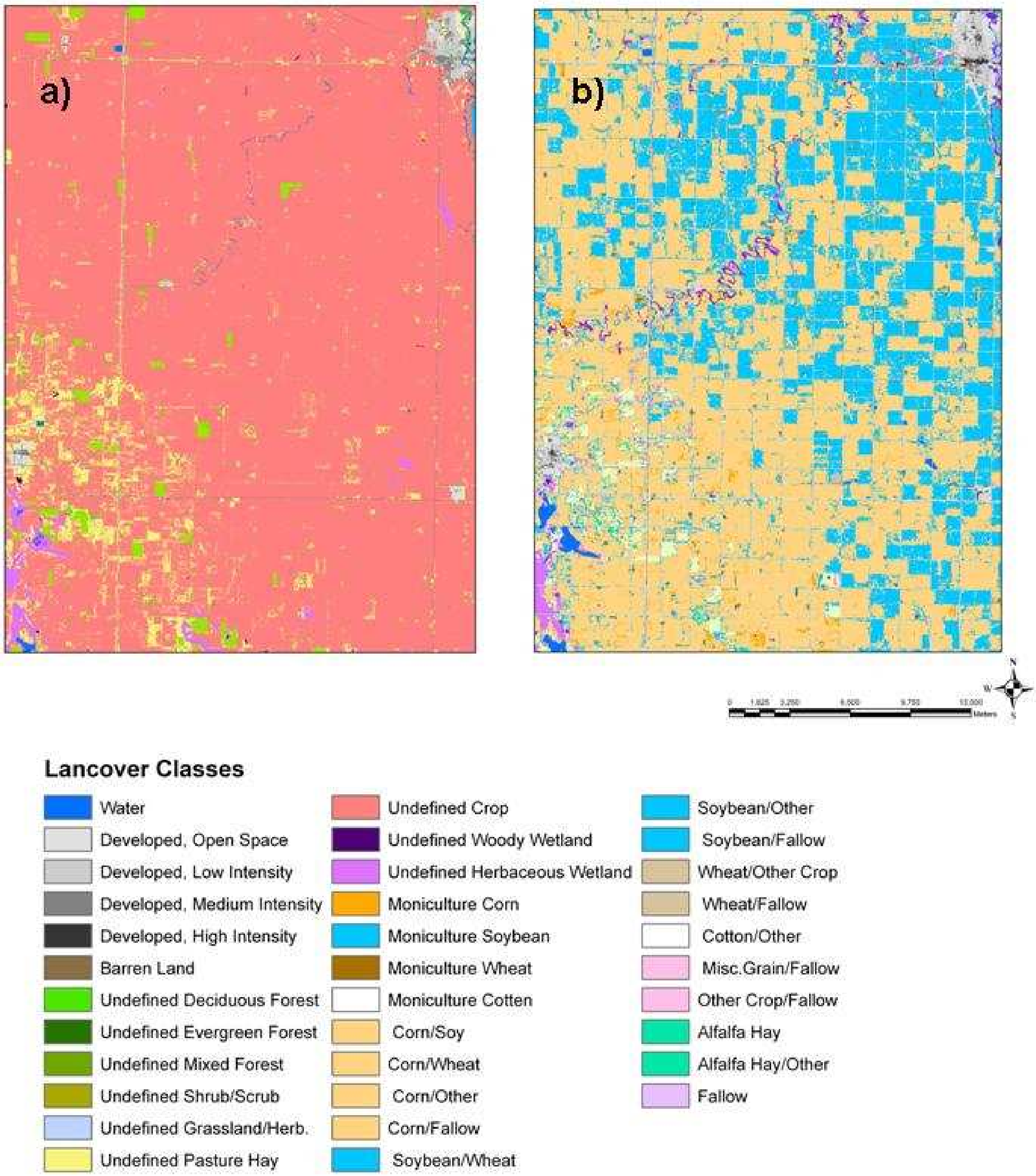
Table 4. Comparison of natural vegetation grid cell counts between the NLCD 2001 and the Midwest classification.

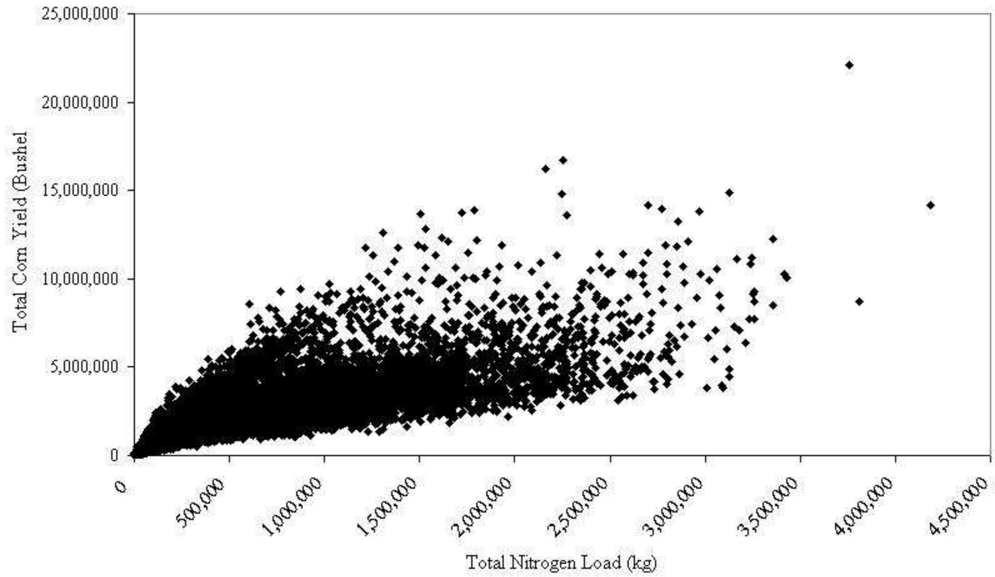
	Midwest Classification								
	Bareground	Deciduous Forest	Evergreen Forest	Mixed Forest	Shrub/ Scrub	Grass-land	Woody Wetland	Herbaceous Wetland	Percent Agreement
NLCD 2001 Classification	Pixel Counts x 1000								%
Bareground	5,151	0	0	0	0	49	0	0	99.05
Deciduous Forest	0	274,973	3,707	31,606	374	9,532	25,160	2,774	78.99
Evergreen Forest	0	7,760	13,335	20,908	266	1,237	5,616	294	26.99
Mixed Forest	0	10,156	1,165	5,251	22	347	1,070	118	28.96
Shrub/ Shrub	0	4,570	1,101	1,769	5,653	13,330	1,660	540	19.75
Grassland	0	8,383	1,645	2,661	34,067	411,600	6,068	7,228	87.27
Woody Wetland	0	5,458	1,014	5,432	25	354	63,540	1,789	81.87
Herbaceous Wetland	0	2,927	208	647	47	2,760	11,790	22,931	55.51
Percent Agreement	100.00	87.51	60.14	7.69	13.97	93.71	55.30	64.28	

Overall Agreement = (802433 x1000 / 1040067 x 1000) = 77%


N
W
E
S
0 200 400 800 1,200 1,600
Kilometers


LAND COVER
NLCD01
CDL
LANDFIRE
Raster
Object ID
GRID value
LANDCOVER
Attribute Table
GRID value
State FIPS
County FIPS
NLCD01
NLCD Class
NLCD row crop (yes, no)
CDL
CDL Value (2004-2007)
Rotation Combination
Final Rotation Class
LANDFIRE
LANDFIRE Value
LANDFIRE Class
NASS
County/District Polygons
County FIPS
District Code
Yield Table
County FIPS
District Code
Crop Yield
Irrigated
Non-Irrigated
Total
ARMS
State Polygons
State FIPS
Crop Management Table
State FIPS
Crop Name
Residue %
Tillage
Weed Control
Pesticide
Insecticide
Herbicide
Fungicide
MIDWEST DATASET
Attribute Table
GRID value (15.5 M +)
LANDCOVER (3 variables)
ARMS (6 variables)
SSURGO (2 variables per crop)
NASS (3 variables per crop)
IRRIGATION (1 variable)
SSURGO
Soil map unit Polygons
MUKEY
COKEY
Yield Table
MUKEY
COKEY
Crop Yield
Irrigated
Non-Irrigated
IRRIGATION
Irrigated Raster
GRID value
Irrigation Attribute Table
GRID value
Percent of land Irrigated (0-100)


a)
b)
Lancover Classes
Water
Developed, Open Space
Developed, Low Intensity
Developed, Medium Intensity
Developed, High Intensity
Barren Land
Undefined Deciduous Forest
Undefined Evergreen Forest
Undefined Mixed Forest
Undefined Shrub/Scrub
Undefined Grassland/Herb.
Undefined Pasture Hay
Undefined Crop
Undefined Woody Wetland
Undefined Herbaceous Wetland
Moniculture Corn
Moniculture Soybean
Moniculture Wheat
Moniculture Cotten
Corn/Soy
Corn/Wheat
Corn/Other
Corn/Fallow
Soybean/Wheat
Soybean/Other
Soybean/Fallow
Wheat/Other Crop
Wheat/Fallow
Cotton/Other
Misc.Grain/Fallow
Other Crop/Fallow
Alfalfa Hay
Alfalfa Hay/Other
Fallow


25,000,000
20,000,000
15,000,000
10,000,000
5,000,000
0
Total Corn Yield (Bushel
0
500,000
1,000,000
1,500,000
2,000,000
2,500,000
3,000,000
3,500,000
4,000,000
4,500,000
Total Nitrogen Load (kg)