The USDA NASS Cropland Data Layer Program

Transition from Research to Operations

(2006-2009)

Claire G. Boryan, Geographer
United States Department of Agriculture
National Agricultural Statistics Service
May 10, 2010

1. INTRODUCTION

The mission of the National Agricultural Statistics Service (NASS), an agency of the United States Department of Agriculture (USDA) is "*to provide timely, accurate and useful statistics in service to U.S. agriculture*". Towards this goal, NASS conducts hundreds of surveys every year collecting information on virtually every aspect of agricultural activity. In 2009, the NASS Cropland Data Layer Program played an important role toward fulfilling this mission using remote sensing techniques to provide operational in-season acreage estimates to the NASS Agricultural Statistics Board (ASB) and Field Offices (FOs) for twenty-seven states and fifteen crops.

The Cropland Data Layer product (Fig. 1) is a comprehensive, raster, geo-referenced, crop-specific land cover classification with a ground resolution of 56 meters, which utilizes satellite imagery to accurately locate and identify field crops. For the first time in 2009, these freely available crop-specific Geographic Information System (GIS) data layers were created for all forty-eight contiguous United States. These GIS products are valuable resources for government agencies, private sector organizations, scientists, educators and students that use land cover information for environmental, agricultural, business or research purposes.

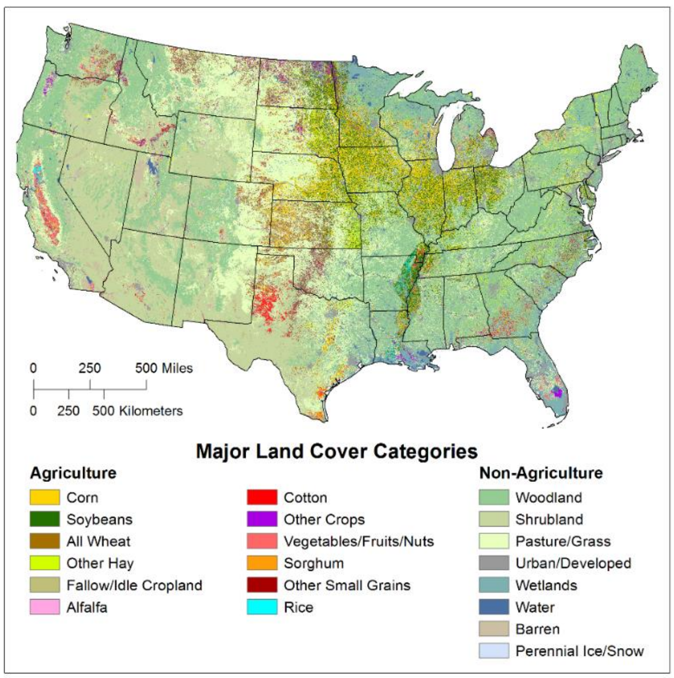

Figure 1. The USDA NASS Cropland Data Layer

The Cropland Data Layer products have been used in a variety of research applications which include assessing the utility of 500 meter (m) Moderate Resolution Imaging Spectroradiometer (MODIS) Time-Series Data for mapping corn and soybean in the U.S. (Chang et al., 2007), validating plant functional type maps developed from MODIS data using multisource evidential reasoning (Sun, et al., 2008), examining the relationship between agricultural chemical exposure and cancer (Maxwell, et al., 2010) and flood mapping assessment with satellite images (Shan, et al., 2010). The Cropland Data Layer was also used to assess the utility of using high resolution aerial imagery to monitor tree cover in agricultural landscapes in North and South Dakota (Liknes, et al., 2010) and to assess automated determination of management units for precision soil conservation (Gelder, et al., 2008).

2. BACKGROUND

NASS initiated its remote sensing acreage estimation program, in the 1970s and early 1980s, with the Large Area Crop Inventory Experiment (LACIE) and Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing (AgRISTARS). The objective of these programs was to determine if crop acreage estimates could be derived using multispectral imagery and ground reference data. These programs were successful at generating unbiased statistical estimates of crop area at the state and county level and, more importantly, reducing the statistical variance of acreage indications from farmer reported surveys (Craig, 2009). The NASS remote sensing acreage estimation program evolved over the years, paving the way for the current Cropland Data Layer (CDL) program, which has been in existence since 1997.

Cropland Data Layer image products and acreage estimates were originally produced using Landsat imagery, NASS survey data, and an in-house remote sensing and estimation software known as Peditor. NASS and the University of Illinois Center for Advanced Computing developed a customized program called Editor. It was transferred to other computer platforms by NASS and the name was modified to Peditor. The historic Peditor method delivered state- and county-level indications in late December for the Crop Production Annual Summary (Craig, 2009).

Beginning in 2006, the CDL program underwent a major restructuring and modernization effort. The original software and data inputs were replaced with a commercial suite of software including Rulequest Research's See5 decision tree software, ERDAS Imagine 9.1 remote sensing software, Environmental Systems Research Institute's (ESRI) ArcGIS, Statistical Analysis Software, Resourcesat-1 Advanced Wide Field Sensor (AWiFS) data, and Common Land Unit data from the Farm Service Agency (FSA). The tremendous efficiency gains resulting from the modernization of the CDL program allowed for the generation of in-season crop acreage estimates, which was previously an unattainable goal.

This paper describes the transition of the CDL program from research to operations, providing in-season acreage estimates in 2009 to meet NASS production deadlines for twenty-seven states and fifteen crops multiple times during the growing season. This overview will describe the original CDL method; the factors leading to change**;** the emergence and use of administrative data as a source of ground reference training and validation data; the shift from Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) to Resourcesat-1 for satellite imagery; the current method and software; the scope, timing and delivery of products; accuracy; and CDL image product delivery.

3. FACTORS LEADING TO CHANGE

The major drawback of the original CDL method, which used Peditor and Remote Sensing Program software, Landsat data and NASS's June Area Survey (JAS) segment data as ground reference data, was the time required to generate the CDL image products and acreage estimates. From 1997-2006, estimates were delivered to NASS headquarters by mid-December of each year for consideration in setting the official county estimates for major "Program Crops", which are those crops for which famers receive subsidies. Over these years, the number of states and crops included in the remote sensing estimation program grew but crop acreage estimates could not be obtained by October 1, which is the deadline for consideration in the October Crop Production Report. Nonetheless, the CDL acreage estimates were used at the state level as supplementary indications, after the fact, and at the county level for operational purposes. By making changes to the CDL program to increase the efficiency, in-season estimates of major crops could be provided to the NASS Agricultural Statistics Board and the state field offices. Furthermore, a still larger goal of the CDL program was to not only provide state-level acreage estimates to the Agricultural Statistics Board in October, but to also meet all estimation deadlines beginning with the June Crop Acreage Report through the December Crop Production Report.

4. GROUND REFERENCE DATA

Both NASS and FSA collect field-level crop information. NASS collects the JAS segment data and FSA collects Common Land Unit polygon data as illustrated in Fig. 2. The scope of the FSA Common Land Unit (CLU) program is comprehensive including all states and extensive coverage of major crops. The program is run at the county level in over 2,300 FSA county offices. There are two important differences between using NASS JAS and FSA CLU data as ground reference training data in the CDL program. First, the individual polygon boundaries of the JAS segments are regularly digitized to support the survey, but the individual fields within each segment would require additional digitizing, a time consuming and labor intensive process. The FSA CLU polygon data, on the other hand, are manually digitized and crop-specific attribute data are collected in the FSA county offices, as part of a standardized program that collects information on all fields included in the FSA programs for compliance and administration purposes (Mueller et al., 2009; Heald, J., 2002). A second difference is that the coverage of major crops provided by FSA is more comprehensive than the JAS segments. In fact, the FSA CLU data approximates full coverage in large production states

There are several shortcomings to using the FSA data. First, a relatively large proportion of CLU polygons include more than one crop type, while JAS segments are digitized to the field with only one crop represented (Craig, 2005). In order to use the FSA data, CLUs with mixed crop types, except certain double crops such as winter wheat followed by soybeans, are excluded from the ground reference data. Second, specialty crops are not well represented in the FSA data leading to a bias toward major speculative crops, which are those for which farmers received subsidies. Third, farmers do not report the crops grown in all CLU polygons each year (Craig, 2001; Mueller et al., 2009).

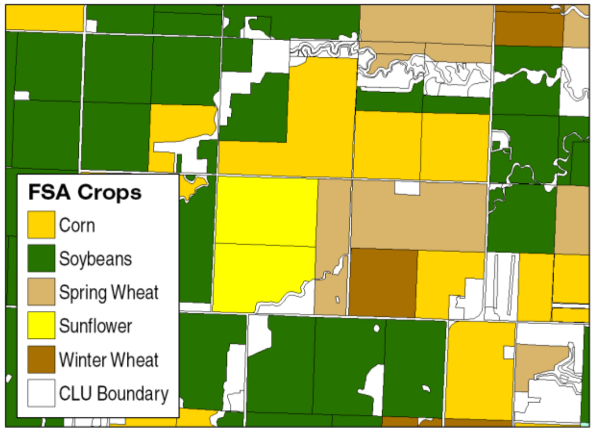

Figure 2. FSA CLU polygon data

Fortunately, these shortcomings are overshadowed by the sheer volume of crop data available from the FSA CLU program. Being a comprehensive agricultural data set that requires minimal preparation and can be updated multiple times, during the growing season, to include the most accurate and recent farmer reported data, greatly outweighs the disadvantages. Using the FSA CLU and 578 attribute data for training has dramatically increased the volume and timeliness of available ground reference data and thereby increased the scope, efficiency and accuracy of the operational CDL program.

5. SATELLITE IMAGERY

In the late 1990s, NASS used both Landsat TM and ETM+ data with a 30 meter spatial resolution in CDL production. The Landsat sensors have a 185 km swath; seven spectral bands including a visible blue, visible green, visible red, near infrared red (NIR), two mid infrared (MIR) bands and a thermal band; a 16-day repeat and eight bit quantization. The synchronization of the two sensors to achieve an eight-day repeat cycle was appropriate for acquiring crop information during the growing season. Landsat data were purchased and made available to NASS via the USDA's Foreign Agricultural Service (FAS), which established the satellite image archive for the purpose of coordinated purchases of satellite imagery for the entire Department of Agriculture (Craig, 2009).

On May 31, 2003, the Landsat 7 ETM+ sensor experienced an anomaly in its scan line corrector. At the time, the imagery was considered unusable by NASS and the CDL program experienced a 50% reduction in the inventory of available satellite imagery. In 2004, the USDA purchased imagery, for evaluation purposes, from the Indian Remote Sensing Satellite (IRS)

RESOURCESAT-1 launched in October of 2003. The moderate spatial resolution (56 meter) Advanced Wide Field Sensor (AWiFS) data were selected for evaluation as a substitute for Landsat imagery in CDL production. The AWiFS imagery was found effective for crop acreage estimation (Boryan and Craig, 2005; NASS, 2006; Seffrin, 2007; Johnson, 2008) after which time NASS decided along with its partner, the FAS, to purchase AWiFS data exclusively for the USDA's satellite imagery archive.

In 2006, NASS began using AWiFS data as the primary source of imagery. The AWiFS sensor offers a moderate spatial resolution (56 meter), a large swath width (720 km), appropriate spectral characteristics for agriculture monitoring and a rapid revisit (5-day repeat) capability. The 56-meter spatial resolution, though coarser than Landsat's 30 meters is sufficient for the accurate identification of large homogenous crop fields (NASS, 2006). Additionally, the full swath width of 720 kilometers, when using both camera A and B acquisitions, provides an excellent opportunity for large area coverage with single day acquisitions. AWiFS offers four spectral bands that closely resemble the most useful of Landsat 5 TM and Landsat 7 ETM+. The sensor acquires data in the visible green, visible red, NIR and short wave infrared (SWIR) bands. The 5-day temporal resolution of AWiFS is a significant improvement from the 16-day revisit of Landsat 5 TM providing the opportunity for abundant nearly cloud free imagery collected throughout the growing season.

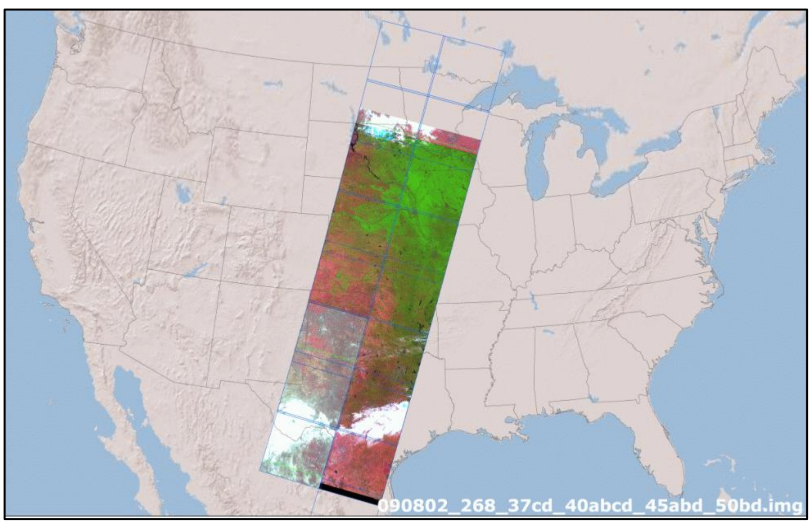

Figure 3. IRS Resourcesat 1 – Advanced Wide Field Sensor Imagery acquired on August 2, 2009. Acquisition descriptions include path/row/quad information. The brightly colored quads are those used in CDL processing.

From the 2006 - 2008, AWiFS imagery was collected from April 1 through the month of October, which is the summer growing season. Acquisitions were excluded based on a 50% cloud cover criteria. With the transition to See5 Decision Tree Software, NASS was able to use a large volume of satellite imagery spanning multiple dates to create classifications. This expansion in the number of scenes acquired throughout the growing season allowed for better separation of crop types based on the varying crop phenological cycles. In 2006, Moderate Resolution Imaging Spectroradiometer Data (MODIS) 16-day Normalized Difference Vegetative Index (NDVI) composites began to be used in the classification process. With its 250-meter spatial resolution, MODIS data could not replace AWiFS data but was useful when collected during the late fall over specific states where the winter wheat crop was beginning to emerge. These fall collections were used to estimate winter wheat acreage in large producing states to support JAS estimation.

In 2009, NASS regularly supplemented AWiFS data with Level 1T (terrain corrected) Landsat 5 TM and Landsat 7 ETM+ data for CDL production, as the entire USGS Landsat Data Archive became available for public consumption at no charge (USGS, 2010). The Landsat data were downloaded from Glovis (https://glovis.usgs.gov). Post processing steps included converting the data from GeoTIFF to ERDAS Imagine image (.img) format, reprojecting from Universal Transverse Mercator to Albers Conical Equal Area, resampling from 30 to 56 meters using cubic convolution resampling method, and mosaicing same day acquisitions.

During the 2009 CDL season, AWiFS experienced technical problems including an on-board data recorder failure and degraded solar panel capacity. Furthermore, increased competition from international customers reduced the availability of AWiFS data for purchase over the United States by the Foreign Agricultural Service Satellite Imagery Archive. Fortunately, the freely available Landsat data could be used as a source of supplemental imagery. The CDL program would not have been able to meet all program deadlines and expand its scope to include the forty eight conterminous states without the use of Landsat data.

6. REMOTE SENSING CLASSIFICATION SOFTWARE

In 2004, transitioning the CDL program from research to operational status appeared to be in the realm of possibility. Changes including new imagery, ground reference data, and image processing and estimation software were required. Already in place were the FSA CLU data, which provided an expansive source of agricultural ground reference data and required no in-house digitization, a significant advance. Additionally, the JAS segment boundaries could still be used as an independent data source for regression modeling. Also available were the AWiFS data, which showed promise for large area coverage at a 5-day repeat cycle. The next step was the identification of commercial remote sensing software that could perform the functions of the Peditor maximum likelihood classifier software.

NASS evaluated ERDAS Imagine, Definiens' eCognition and Rulequest Research's See5 decision tree software. The remote sensing software selected needed to be affordable, efficient and accurate. See5 came highly recommended by EROS Data Center researchers and was used to produce the National Land Cover Database for 2001 and through literature reviews (Homer, C. et al., 2004; 2007; Hansen et al., 1996; Friedl and Brodley, 1997 and Lawrence et al., 2004) was found to be the most appropriate replacement for the Peditor maximum likelihood classifier.

See5 was the remote sensing classification software used by NASS from 2006 -2009 and was the primary driver of the expansion of the CDL program. The most significant difference between Peditor and See5 was the time required to produce a statewide CDL. Once the See5 method was fully developed, an experienced analyst could produce a statewide CDL, after all preprocessing of ground reference data, imagery, and ancillary data was complete, within several days. It required between one to two months for the same analyst, using the Peditor method, to produce a statewide CDL product. The difference was in large part because See5 was able to generate a statewide CDL in one process incorporating all input data.

Although Peditor was an excellent classifier, a number of limitations made the classification process more time consuming. Peditor operated by creating multiple smaller classifications. The intersection of same date Landsat scenes defined "analysis districts". A separate classification was generated for each analysis district. Using the Peditor method, some states required as many as twelve separate analysis districts, which in turn required running twelve separate classifications to produce a statewide CDL. The individual classifications were merged to create the statewide CDL mosaic.

With See5, even though by definition it classifies the intersection of inputs, there is a technique to get around this obstacle so that the entire state or region can be classified in one process. All input data including imagery and ancillary data must be set to a specific map extent when created. Consequently, even though all of the imagery does not cover an entire state, if all of the inputs are set to this specific map extent, then See5 categorizes all land cover within this region or state. While it takes additional time preparing the input data, the amount of time saved in the classification phase is significant.

Additionally, See5 provides options that improve the quality and accuracy of the CDL products. These options include allowing for the ingestion of an abundance of satellite imagery and other non-parametric data sources; incorporating a boosting algorithm in which the classifier reviews the results multiple times to refine or "prune" the decision tree; tolerating image noise, such as clouds, haze or even gaps in the imagery and generating confidence layers that corresponded to the resulting classifications. Lastly, the National Land Cover Database (NLCD) 2001 can be used with See5 for training on non-agricultural categories and can be combined with the agricultural training to create a complete training set for the state or region.

In 2009, NASS used AWiFS, MODIS, Landsat TM and ETM+ data to produce the CDL products. Imagery was acquired from the fall of 2008 until late September 2009. Using imagery collected over the entire growing season facilitated the separation of crop phenologies and the accurate identification of cropland. In some instances, over a particular area, six or more satellite scenes acquired during the growing season were used to classify the land cover. This was extremely useful when attempting to identify double crops such as winter wheat followed by soybeans or crops with similar phenologies. Peditor could only ingest a maximum of two scenes of a study area, a significant limitation.

Nonparametric data sets such as the United States Geological Survey (USGS) Digital Elevation Model, USGS percent canopy layer, and USGS percent impervious layer were used from 2007 – 2009 to help identify non- agricultural categories and separate them from crops. The USGS Digital Elevation Model is most useful in regions with significant topographic variation. Further, crops are most often grown in areas of low topographic relief. For example, in Mississippi, Louisiana and Missouri, a significant percentage of the agriculture grown in the region is located in the low lying portion of the Delta. The percent canopy layers help identify the forested areas and the percent impervious layers helps identify urban infrastructure. These raster layers could not be used with Peditor.

Boosting or bagging, in which the classifier reviews the results multiple times to refine or "prune" the decision tree, was available with See5. This was shown to improve accuracy in the literature (Quinlan, J., 1996). In 2009, ten boosts were generally run to refine the CDL classification. Boosting was not available with Peditor.

The NLCD 2001 is currently used for training for non-agricultural categories. The NLCD 2001 was released in 2006 at which time NASS began using it for non-agricultural sampling. When using Peditor, an analyst had to manually create non-agricultural ground reference training data. "Extra signatures" were created for clouds, water, grass, trees, wetlands and many other non-agricultural categories, a very labor-intensive process. Additionally, these "extra signatures' were created for each individual classification or analysis district with Peditor.

A tremendous advantage of See5 and improvement in operational efficiency was the software's tolerance of image noise such as clouds, haze and the scan gaps in the Landsat 7 ETM+ data. As long as there was an abundance of clear imagery overlaying the same location as the image noise, the software seemingly ignored the bad data. When using Peditor, "extra signatures" had to be created for all analysis districts in which clouds were evident.

7. REMOTE SENSING PROGRAM SOFTWARE TO ESRI ARCGIS

Starting in 2006, when the FSA CLU data became the primary source of ground reference data for the CDL program, the switch was made from Remote Sensing Program (RSP) software to ESRI's ArcGIS software. ArcGIS was the clear choice as USDA had an enterprise software license and many staff were trained in its use. The preparation of the FSA CLU data was dramatically more efficient when using ArcGIS than it was with RSP software.

Models were written in ArcGIS to merge the original county FSA CLU shape files into statewide shape files. The shape files were then cleaned, reprojected to Albers Conical Equal Area projection and buffered inward 56.0 meters. All of these steps, which were relatively time consuming for 48 states, were completed on the CLU polygon data in 2009, prior to the start of the growing season. All of these processes could not be performed with RSP, which was primarily used for digitizing and editing crop attribute information. During the crop season, the JAS segment data required approximately one month for digitizing in the NASS Field Offices and two weeks for editing by an analyst.

Once the FSA CLU polygons were linked to the FSA 578 attribute data, ArcGIS models were used to exclude non-matching CLUs, separate CLUs into training and validation data sets, and rasterize the shape files for use in See5. The ArcGIS models dramatically improved the efficiency of the process whereby the most current ground reference training data could be used prior to in-season deadlines. ESRI's ArcGIS was an important contributor improving the efficiency and quantity of the ground reference data available for use in producing the CDLs.

8. PEDITOR TO STATISTICAL ANALYSIS SOFTWARE

From 1997 to 2005, Peditor performed all of the functions of both remote sensing classification and estimation software system. Once the decision was made to transition from Peditor to See5, new estimation software was required. Statistical Analysis Software (SAS) was selected as it was widely used within NASS and had the statistical analysis capabilities that NASS required. In 2006, the regression estimator in Peditor was well developed and documented (Day, 2002). Consequently, the identical programs written and run in Peditor were transitioned to SAS.

By 2007, SAS was able to increase the efficiency of estimation modeling and help transition the CDL program from research to operational status. One of the most important advantages of SAS was the ability to interactively review the regression analysis results in IML Workshop (called Stat Studio in SAS 9.2). This made the process of removing outliers and rerunning the regression modeling less labor intensive for statisticians. Second, the original format of JAS segment data was formatted in SAS and made the data easier to use. Third, result tables in SAS were output in PDF files and Excel files, which were easier for NASS Headquarters statisticians to import and analyze than the ASCII` tables generated in Peditor. Another important advantage that occurred during the transition from Peditor to See5 and SAS was the ability to run estimates for the entire state at one time.

9. SCOPE AND TIMING

NASS's Cropland Data Layer program has been in existence since 1997. For the first 10 years, Peditor was the software of choice as no other existed that could run remote sensing classifications and regression modeling. During this ten- year period, the program expanded from two states in 1997 to nineteen states in 2006 delivering state and county estimates to the NASS Agricultural Statistics Board in mid-December of each year for consideration in setting county-level estimates. With the transition of the program to a commercial software suite; including See5, ArcGIS, ERDAS Imagine, and SAS; tremendous technological and methodological efficiencies were achieved moving the program from research to operational status. In 2007, the CDL became operational providing acreage estimates for thirteen states and nine crops to the NASS Agricultural Statistics Board for the October Crop Report. For the first time, remote sensing estimates were used in consideration for setting the NASS official state acreage estimates, a milestone for the program. An additional eight CDL state level image products were generated in the post-season for a total of twenty-one 2007 CDL products.

To continue to enhance the scope and timing of the program, research was conducted to determine whether accurate estimates could be provided earlier in the season (Boryan et al., 2008). Acreage estimates provided to the NASS Agricultural Statistics Board to meet the June Acreage, August, September, and October Crop Production Reports, as well as, the September Small Grains Report would further NASS's remote sensing mission.

In 2009, the CDL program expanded to provide estimates to the Agricultural Statistics Board meeting all production deadlines for twenty-seven states and fifteen crops. Twenty additional CDL image products were generated after the growing season for a total of forty-eight 2009 statewide CDL products. For the first time, CDL image products were created for all conterminous states in the U. S.

10. ACCURACY

The accuracy of the CDL agricultural crop categories were calculated by comparing the CDL image products with independent validation files from the FSA CLU data. During the ground reference preparation phase, 30% of the available FSA CLU polygons data were set aside for use in accuracy assessment. The producer and user accuracies were generally 85% to 95% correct for the major crop-specific land cover categories. The producer's accuracy relates to the probability that a FSA CLU ground reference pixel was correctly mapped and measures errors of omission. The user's accuracy indicates the probability that a pixel from the classification actually matches the FSA ground reference data and measures the errors of omission (Congalton and Green, 1999). Accuracies for the non-agricultural categories were not provided as these categories were not based on known ground reference data, but on a separate land cover classification, the 2001 NLCD, with its own inherent rates of error.

11. CROPLAND DATA LAYER IMAGE PRODUCT DELIVERY

In May, 2010, the CDL image products were downloadable free of charge from the National Resource Conservation Service (NRCS) Geospatial Data Gateway at https://datagateway.nrcs.usda.gov. The CDL image products were raster, georeferenced, categorized land cover data layers with a spatial resolution of 30 meters (CDLs produced prior to 2006 using TM data) and 56 meters (CDLs produced from 2006 - 2009 using AWiFS data). The CDL products on the Geospatial Data Gateway were provided in GeoTIFF format, Universal Transverse Mercator (UTM), North American Datum (NAD) 1983 or World Geodetic System 1984 map projection. The CDL products were available on CD as GeoTiff (.tiff) or ERDAS imagine (.img) in two map projections including UTM and Albers, NAD 1983, Geodetic Reference System 1980**.** The CDLs were aggregated to standardized categories emphasizing agricultural land cover. The associated metadata for each CDL product were included with the Geospatial Data Gateway download.

12. CONCLUSION

This paper describes how the USDA NASS Cropland Data Layer program transitioned from a research to operational program by means of the production efficiencies achieved using the FSA CLU data, AWiFS Imagery, See5, ArcGIS and SAS software. In 2009, the NASS Cropland Data Layer program provided operational in-season acreage estimates to the NASS Agricultural Statistics Board and Field Offices for twenty-eight states and fifteen crops. The CDL program covered all NASS speculative program crops providing updated acreage estimates throughout the growing season using the most up–to-date farmer-reported and satellite data available. Additionally, for the first time in 2009, the freely available CDL products were created for the forty-eight conterminous states in the U.S.

Having achieved this level of coverage in 2009, it is the goal of the CDL program to continually provide yearly updates at the state level for research and operational applications, including climate change, forest fire assessment, crop rotation and watershed analysis, agricultural yield, assessments of urban expansion and much more. Further, the CDL program will continue to evaluate its ability to expand the quantity and scope of crop acreage estimates provided to the NASS Agricultural Statistics Board and Field Offices to further the NASS mission of providing the most timely, accurate and useful agricultural statistics possible.

13. REFERENCES

Boryan, C., and M. Craig (2005). Multiresolution Landsat TM and AWiFS Sensor Assessment for Crop Area Estimation in Nebraska, *Proceedings from Pecora 16*, Sioux Falls, South Dakota.

Boryan, C., M. Craig, M. Lindsey (2008). Deriving Essential Dates of AWiFS and MODIS for the Identification of Corn and Soybean Fields in the U.S. Heartland, *Proceedings from Pecora 17*, Denver Colorado, November 2008.

Chang, J. C., M. C. Hansen, K. Pittman, M. Carroll and C. DiMiceli (2007). "Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets," *Agronomy Journal*, 99: 2007, 1654-1664.

Congalton, R. G., and K. Green (1999). *Assessing the accuracy of remotely sensed data: principles and practices*. Boca Raton: Lewis Publishers.

Craig, M. (2001). The NASS Cropland Data Layer Program. Presented at the Third International Conference on Geospatial Information in Agriculture and Forestry, Denver, Colorado, 5-7 November 2001.

Craig, M. (2005). "Using FSA Administrative Data in the NASS Cropland Data Layer" Draft as of 9/7/2005; write-up of FSA data used for Nebraska 2002-2004 research; circulated administratively only in NASS; NASS/RDD/GIB/SARS Fairfax, VA 2005.

Craig, M. (2009). A Brief History of the Cropland Data Layer at NASS https://www.nass.usda.gov/Research_and_Science/Cropland/CDL_History_MEC.pdf Accessed May 10, 2010

Day, C.D. (2002) "A Compilation of PEDITOR Estimation Formulas". RDD Research Paper RDD-02-03, USDA, NASS, Washington, D.C. January, 2002.

Friedl, M., and C. Brodley (1997). Decision Tree of Land Cover from Remotely Sensed Data. *Remote Sensing of the Environment*, 61: 339-409.

Gelder, B., R. Cruse, and A Kaleita (2008). "Automated determination of management units for precision conservation," *Journal of Soil and Water Conservation*, September 2008 vol. 63 no. 5 pp. 273-279.

Hansen, M., R. Dubayah, & R. Defried. (1996) Classification Trees: An Alternative to Traditional Landcover Classifiers. *International Journal of Remote Sensing*, 17, 1075-1081.

Heald, J (2002). "USDA establishes a Common Land Unit, "ESRI ArcUser Online. April-June 2002, <http://www.esri.com/news/arcuser/0402/usda.html > Accessed, February 10, 2010.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan, (2004). "Development of a 2001, National Land Cover Database for the United States", *Photogrammetric Engineering & Remote Sensing*. Vol. 70, No. 7, pp 829-840, July 2004.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. Van Driel, and J. Wickham (2007). "Completion of the 2001 National Land Cover Database for the Conterminous United States", Photogrammetric *Engineering & Remote Sensing*. Vol. 73, No. 4, pp. 337-341, April 2007.

Johnson, D. M. (2008). A comparison of coincident Landsat-5 TM and Resourcesat-1 AWiFS imagery for classifying croplands, *Photogrammetric Engineering and Remote Sensing*. Vol 74, No 11 November 2008.

Lawrence, R., A. Bunn., S. Powell., and M. Zambon (2004). Classification of Remotely Sensed Imagery using Stochastic Gradient Boosting as a Refinement of Classification Tree Analysis, *Remote Sensing of the Environment*, 90.331-336.

Liknes G., C. Perry, and D. Meneguzzo (2010). "Assessing Tree Cover in Agricultural Landscapes Using High-Resolution Aerial Imagery," *The Journal of Terrestrial Observation*, Vol. 2 No. 1 (Winter 2010), pp. 38-55.

Maxwell, S. K., J. Meliker and P. Goovaerts (2010).” Use of land surface remotely sensed satellite and airborne data for environmental exposure assessment in cancer research,” *Journal of Exposure Science and Environmental Epidemiology* (2010), 20, 176–185

Mueller, R., C. Boryan and R. Seffrin (2009). Data Partnership Synergy: The Cropland Data Layer. August ’09 GMU geoinformatics conference < https://www.infona.pl/resource/bwmeta1.element.ieee-art-000005293489/tab/summary>

NASS, 2006. Assessment of TM and AWiFS imagery for cropland classification: three case studies. https://ipad.fas.usda.gov/pdfs/2006/NASS_AWiFS_TM_CDL_comparison_Final.pdf Accessed, February 10, 2010.

Quinlan, J. (1996). Bagging, Boosting, and C4.5, *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Portland , Oregon (American Association for Artificial Intelligence Press, Menlo Park, California), pp. 725 – 730.

Seffrin, R. (2007). Evaluating the Accuracy of 2005 Multitemporal TM and AWiFS Imagery for Cropland Classification of Nebraska, *Proceedings of the ASPRS 2007 Annual Conference*, Tampa Florida, May 7-11, 2007.

Shan J., E. Hussain, K. Kim, and L. Biehl (2010). “Flood Mapping with Satellite Images and its Web Service,” *Photogrammetric Engineering & Remote Sensing*, February 2010.

Sun, W., S. Liang, G. Xu, H. Fang and R. Dickinson (2008). “Mapping plant functional types from MODIS data using multisource evidential reasoning,” *Remote Sensing of Environment*, Volume 112, Issue 3, 18 March 2008, Pages 1010-1024.

USGS (2010). <https://landsat.usgs.gov/products_data_at_no_charge.php>Accessed May 10, 2010.

USGS Newsroom, (2005). “Orthorectified Landsat Digital Data Now Available from the USGS” <https://www.usgs.gov/news/orthorectified-landsat-digital-data-now-available-usgs >