Remote Sensing Applications in Agriculture at the USDA National Agricultural Statistics Service

Jeffrey T. Bailey
Chief of Geospatial Information Branch
Research and Development Division
United States Department of Agriculture
National Agricultural Statistics Service
3251 Old Lee Highway
Fairfax, VA 22030 (USA)
Jeff_Bailey@nass.usda.gov

Claire G. Boryan
Geographer
Research and Development Division
United States Department of Agriculture
National Agricultural Statistics Service
3251 Old Lee Highway
Fairfax, VA 22030 (USA)
Claire_Boyan@nass.usda.gov

Date: September 20, 2010

ABSTRACT

The United States Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS) has researched and used remote sensing technology for acreage estimation since the early 1970s. Significant advancements in recent years have enabled NASS to transition the use of remote sensing from primarily a research function to performing an integral role in the agency's crop acreage estimation program covering all major crops grown in high producing states in the U.S. This accomplishment was achieved in large part as a result of 1) enhanced data partnerships, 2) improved methodologies, 3) increased availability of commercial software, and 4) improved imagery and ancillary data. With acreage estimation now operational, the Agency is focusing its efforts to transition yield estimation from research to operational, as well. Currently corn and soybean yield estimates for 10 major producing states are provided in-season to the Agricultural Statistic Board (ASB) and the NASS field offices. Although, yield estimates are solid, with more experience and research, improvements are expected. Looking to the future, NASS has begun conducting research to quantitatively measure crop progress and condition using remote sensing and intends to expand the use of remote sensing into mapping soil moisture and improving disaster assessments and monitoring.

Key Words: Cropland Data Layer, Remote Sensing, Satellite Imagery, Agriculture

1.0 BACKGROUND

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 US agriculture*". Towards this goal, NASS conducts hundreds of surveys every year collecting information on virtually every aspect of agricultural activity. In 2010, the NASS Cropland Data Layer (CDL) 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 sixteen crops.

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) to determine if crop acreage estimates could be derived using multispectral imagery and ground truth 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). NASS' remote sensing acreage estimation program evolved over the years paving the way for the current CDL program which has been in existence since 1997.

Originally an in-house remote sensing and estimation software known as Peditor, Landsat satellite imagery, and NASS survey data were used to produce the CDL image products and acreage estimates. 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 estimates 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 (SAS), Resourcesat-1 Advanced Wide Field Sensor (AWiFS) data, and Common Land Unit (CLU) data from the Farm Service Agency (FSA). Tremendous efficiency gains were achieved due to the modernization allowing for the generation of in-season crop acreage estimates, a goal never achieved using the older method.

1.0 Recent Successes in Remote Sensing at NASS

For the first time on January 10, 2010 NASS released a National Cropland Data Layer for 48 conterminous states. The CDL product 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. Geographic Information System (GIS) data layers are valuable resources for government agencies, private sector organizations, scientists, educators and students that use land cover information for environmental, agricultural, business or research.

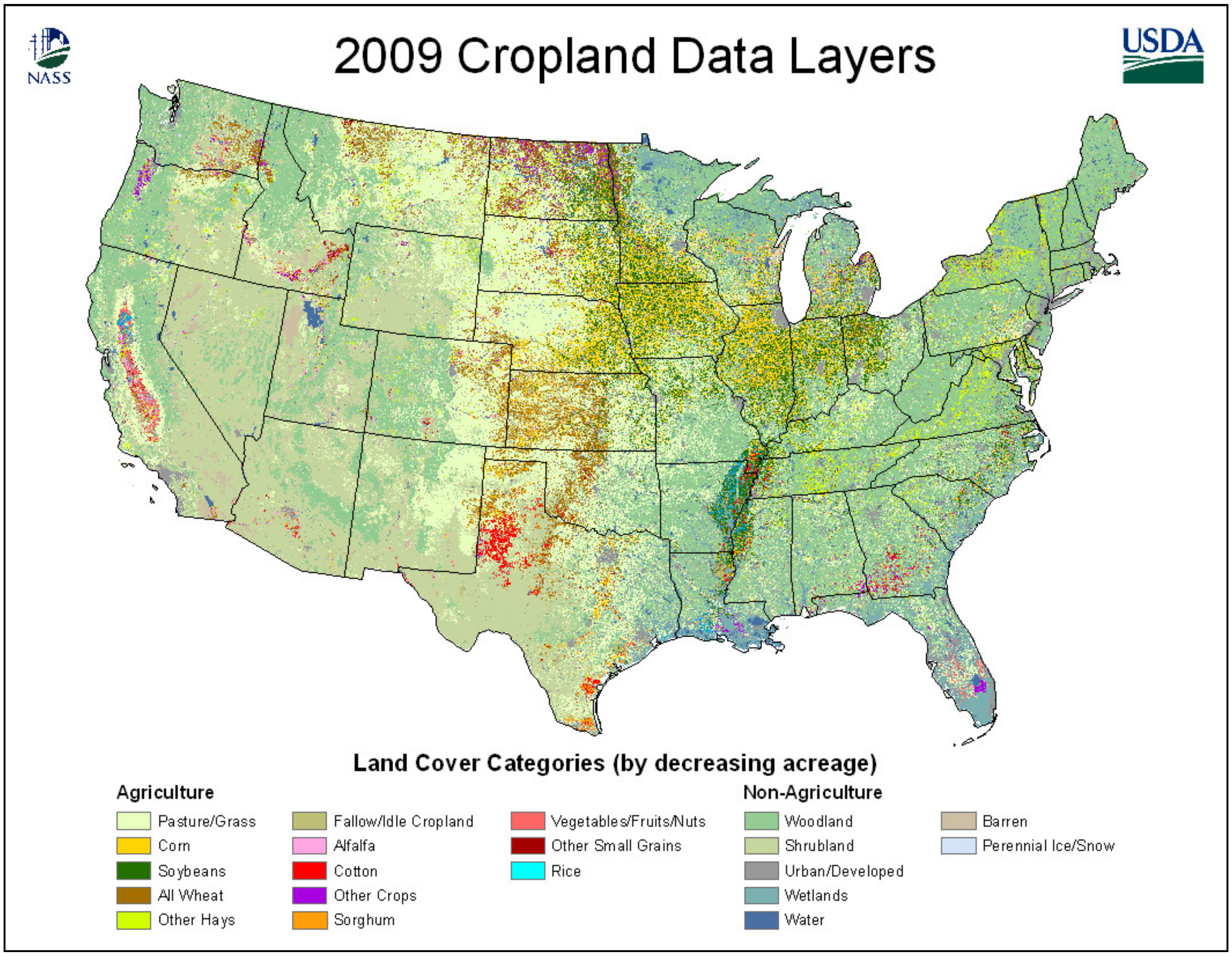

Figure 1. 2009 Cropland Data Layers, Released January 11, 2010.

CDL products have been used in a variety of research applications ranging from assessing the utility of 500m 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), to flood mapping assessment with satellite images (Shan, et al., 2010). The CDL was also used to assess the utility of monitoring tree cover in agricultural landscapes in North and South Dakota using high resolution aerial imagery (Liknes, et al., 2010) and to assess automated determination of management units for precision soil conservation (Gelder, et al., 2008).

Using the CDL as the foundation, NASS runs a regression estimator to produce crop acreage estimates. Beginning in 2007 NASS was able to generate in-season acreage estimates for the first time. Previously, acreage estimates could only be derived at the end of the crop season. Currently, winter wheat acreage estimates are provided to the NASS ASB and the field offices beginning in June. Acreage estimates for corn, soybeans and other crops are generated in August, September and October. In 2007, estimates were provided for six crops in 12 states. The program has expanded to the point that in 2010 NASS produced estimates at the state level for 16 different crops in 27 states. In addition, post season small area estimates will be provided for 36 states.

Again using the CDL as the foundation, NASS provided operational in-season corn and soybean yield estimates to the NASS ASB and field offices for the past two years. In 2009 NASS calculated corn and soybean yield estimates for 6 states. In 2010, NASS anticipates providing yield estimates for 10 states. Additionally at the end of the year, NASS will provide small area yield estimates to be used operationally in the county estimates program. With additional research, NASS anticipates expanding the yield estimation program to include additional states and crops. Currently, the remotely sensed Normalized Difference Vegetative Index (NDVI) is the sole input to the yield models. Additional inputs related to weather and soils are being explored with the anticipation of increasing the accuracy of the yield predictions.

2.0 Critical Components for Remote Sensing

The first critical component for any remote sensing program is solid ground truth information. Without ground data to identify land cover categories, to train the classifier and validate the output image products, it is impossible to run a defensible program that provides reliable results. Ground truth is mentioned first, because it must be seriously considered before initiating plans for any remote sensing application. Secondly, a source of satellite imagery is required. There are many sources of satellite imagery which vary considerably in cost, as well as, spatial, temporal, spectral and radiometric resolution. Finding an imagery source that also provides a guarantee of future continuity is an important consideration, since once a program has been researched and implemented, it becomes more difficult to transition to another satellite. Thirdly, using remotely sensed data requires a sizable investment in Information Technology (IT) resources. However, with the speed of computers continuing to increase and the price of disc storage on the decline this has become much less of a hindrance.

2.1 Ground Truth

NASS has two sources of field level crop information for ground truth, its own June Area Survey (JAS) and the USDA Farm Service Agency (FSA). NASS collects the June Area Survey (JAS) segment data and the FSA collects CLU polygon data. The scope of the FSA 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 JAS and FSA data, as ground truth, in the CDL program. First, the JAS data requires manual digitization of individual segments by NASS staff or cooperators while the FSA data does not. The individual polygon boundaries of the JAS segments are regularly digitized to support the survey but the individual fields within each segment require additional digitization.

The FSA CLU polygon data are digitized and crop specific attribute data collected in the FSA county offices as part of a standardized GIS layer that collects information on all fields in FSA programs on a near real time basis for compliance and administration purposes (Mueller et al., 2009; Heard, J., 2002; Anderson et al., 2005). A second difference is that the coverage of major crops provided by the FSA are more comprehensive than the 150 – 400 one square mile area segments included in the JAS data, approximating full coverage in major speculative states.

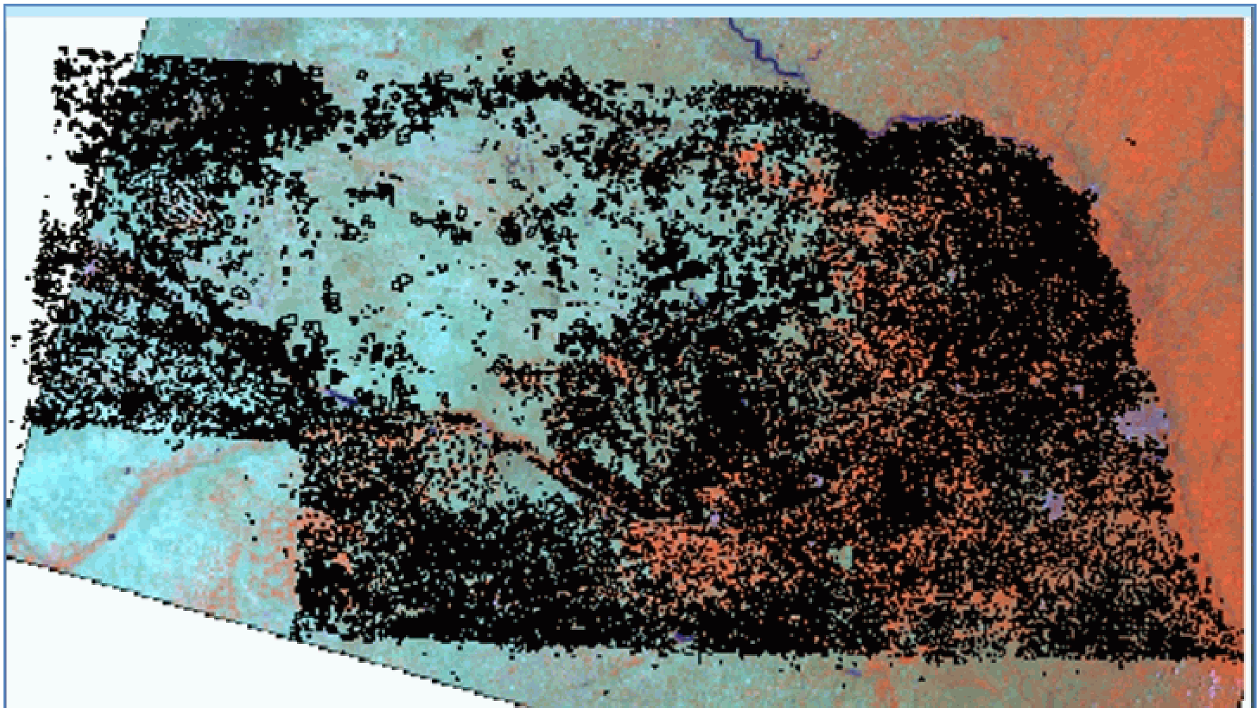

Figure 2. FSA CLU polygon data overlaying an Advanced Wide Field Sensor image

However, there are several shortcomings to using the FSA data. First, approximately fifty percent of CLU polygons include more than one crop type per CLU while JAS segments are digitized to the field (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 truth. Second, specialty crops are not well represented in the FSA data leading to a bias toward "program crops", for which farmers received subsidies. Third, not all CLU polygons are attributed each year (Craig, 2001; Mueller et al., 2009).

Fortunately, these shortcomings are greatly 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 greatly outweighs the disadvantages. Using the FSA CLU and 578 attribute data for training has dramatically increased the volume and timeliness of available ground truth and thereby increased the scope, efficiency and accuracy of the operational CDL program.

2.2 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 8 bit quantization. The synchronization of the two sensors to achieve an 8 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 (SIA) for the purpose of coordinated purchases of satellite imagery for the entire Department of Agriculture (Craig, 2009).

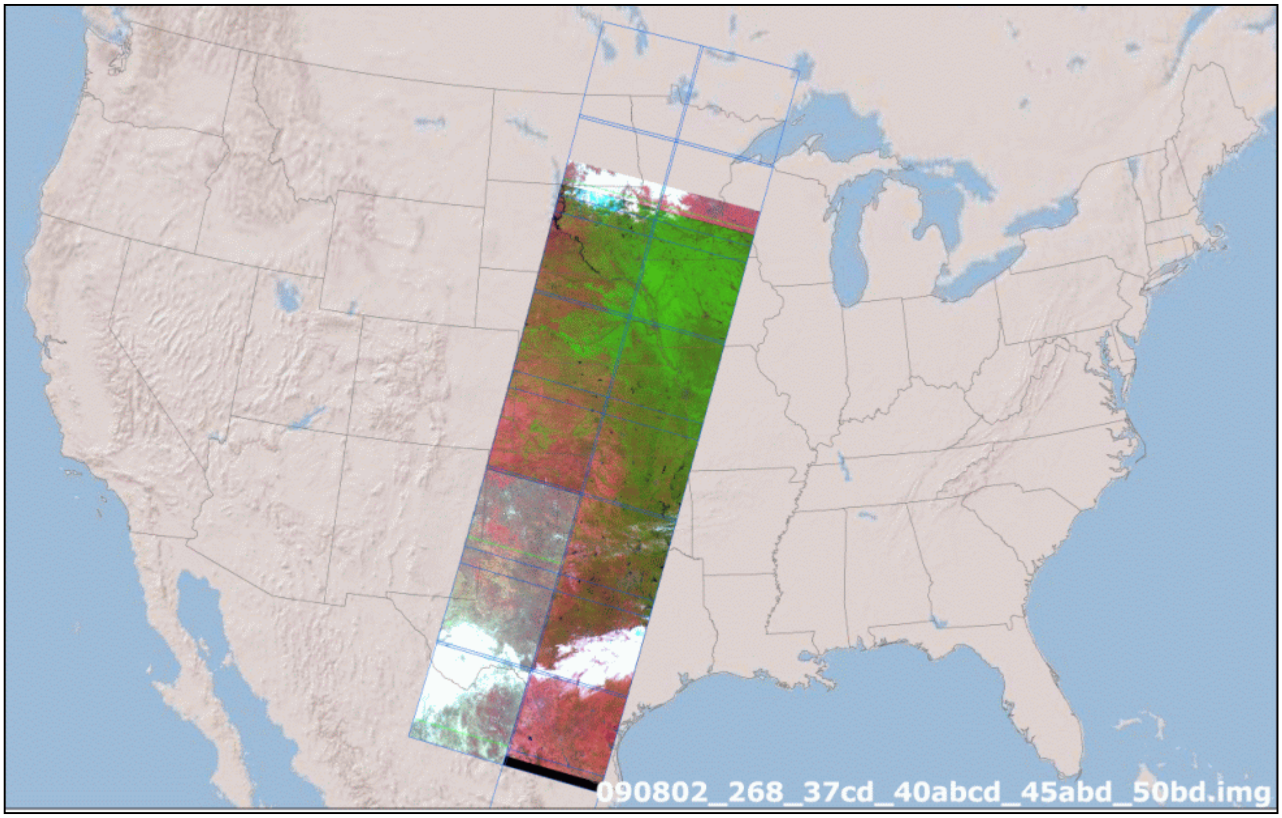

Figure 3. IRS Resourcesat 1 – Advanced Wide Field Sensor (AWiFS) Imagery acquired on August 2, 2009. Acquisition descriptions include path/row/quad information. The brightly colored quads are those used in CDL processing.

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. NASS conducted investigations to assess the effectiveness of AWiFS data for crop acreage estimation including: Nebraska, 2004 (Boryan and Craig, 2005); Arkansas (Delta Region); Nebraska, 2005 and Coincident Studies, (Arkansas, Illinois, Iowa) 2005 (NASS, 2006; Seffrin, 2007; Johnson, 2008) after which time NASS decided along with its partner, FAS, to purchase AWiFS data exclusively for the USDA's SIA, International Productions Assessment Unit.

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, near infrared (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.

From the 2006 - 2008 growing seasons, AWiFS was collected from April 1 through the month of October. Acquisitions were excluded based on a 50% cloud cover criteria. Fortunately, with new software a large volume of satellite imagery and ancillary data could be used in the classification process. 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 could not replace AWiFS but was useful when collected during the late fall over specific states where the winter wheat crop was beginning to emerge.

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, 2005). The Landsat data were downloaded from Glovis (http://glovis.usgs.gov). Post processing steps included converting the data from GeoTIFF to ERDAS Imagine image (.img) format, reprojecting from Universal Transverse Mercator (UTM) to Albers, resampling from 30 to 56 meters using cubic convolution (CC) 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. Further, increased competition from international customers reduced the availability of AWiFS data for purchase over the U.S. by the FAS archive. Fortunately, the freely available Landsat data were available for use as a source of supplemental imagery. The CDL program would not have been able to meet all program deadlines, as well as, expand its scope to include the forty eight conterminous states without the use of Landsat data.

2.3 Software and IT infrastructure

2.3.1 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 truth, image processing and estimation software were required. Already in place was the FSA CLU data which provided an expansive source of agricultural ground truth 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 Peditor, NASS' original in-house remote sensing maximum classifier and estimation 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 (NLCD) 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) and was found to be the most appropriate as a replacement for the Peditor maximum likelihood classifier.

See5 was the remote sensing classification software used by NASS since 2006 and was the primary driver of the expansion of the CDL program. The most important factor was the time required to produce a state wide CDL. Once the See5 method was fully developed, an experienced analyst could produce a state wide CDL, after all pre processing of ground truth, imagery, and ancillary data was complete, within several days. It required one to several months for the same analyst, using the Peditor method, to produce a state wide CDL product. The difference is in large part because See5 is able to generate a state wide CDL in one process incorporating all input data.

Although Peditor was an excellent classifier, there were a number of limitations that made the classification process more time consuming. Peditor operated by creating multiple smaller classifications. The intersection of Landsat scenes defined "analysis districts" (AD). A separate classification would be 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 state wide CDL. The individual classifications were merged to create the state wide CDL.

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, for example the entire state of Nebraska if all of the inputs are set to this specific map extent then See5 categorizes all land cover within this region. This is a tremendous time saver. It takes additional time preparing the input data, but the time saved in the classification phase is significant.

Additionally, See 5 provides options which 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 which corresponded to the resulting classifications. Lastly the 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 throughout 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. This was a significant limitation.

Non parametric data sets such as the USGS Digital Elevation Model (DEM), 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 DEM 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 would have to manually create non agricultural ground truth. "Extra signatures" were created for clouds, water, grass, trees, wetlands and many other non agricultural categories, a very time consuming 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 its 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" would have to be created for all analysis districts ADs in which clouds were evident.

3.3.2 RSP to ESRI ArcGIS

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

Models were written in ArcGIS to merge the original county FSA CLU shape files into state wide shape files. The shape files were then "cleaned, reprojected to Albers Conical Area (Albers) and buffered inward 30.0 – 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 crop season. All of these processes could not be performed with RSP which was primarily used for digitizing and editing crop attribute information. The JAS segment data required approximately one month, during the crop season, for digitizing in the FOs and two weeks for editing by a CDL 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 truth could be used prior to in season deadlines. ESRI's ArcGIS was an important contributor improving the efficiency and quantity of the ground truth available for use in producing CDLs.

3.3.3 Peditor to SAS

From 1997 to 2005, Peditor performed all of the functions of both a remote sensing classification and estimation software. Once the decision was made to transition from Peditor to See5, new estimation software was required. 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, results tables in SAS were output in .pdf files and Excel files that 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.

3.0 Future Research

As previously mentioned, with the crop area estimation program in an operational mode, NASS is currently transitioning the yield estimation work from research to operational, as well. For 2010, NASS expects to model yield estimates in 10 states for corn and soybeans. USDA's Agriculture Research Service (ARS) has been working cooperatively with NASS for the last 10 years to develop yield models using primarily the Normalized Difference Vegetative Index (NDVI.) The results have been good, however there remains room for improvement. NASS is interested in expanding the yield program to cover additional states and crops while expanding the models to include other inputs, such as weather and soils data.

There are four new research areas that NASS hopes to move into within the next year. These research areas include 1) provide quantitative assessments of crop progress, 2) quantitatively assess crop condition, 3) the development of soil moisture data layers that provide quantification of top and subsoil moisture available for crops, and 4) an expansion of natural disaster monitoring and assessments.

4.0 Conclusion.

The NASS Remote Sensing program has been successful in part due to the partnerships that were established over the years. First, the USDA FSA provides the comprehensive Common Land Unit (CLU) ground truth data set. The CLU data cover all the program crops administered by USDA extremely well. Best of all the data are georeferenced and do not require any manual digitization by NASS. A second partnership is the USDA Foreign Agriculture Service (FAS) Satellite Imagery Archive (SIA). The SIA purchases and archives all satellite imagery for all USDA member agencies. This pooling of resources has enabled all members to buy and manage imagery more efficiently. Third, NASS joined the US Geological Survey's (USGS) Multi-Resolution Land Characteristics (MRLC) Consortium in 2006 as a cooperative partner. The MRLC Consortium is a group of federal agencies who first joined together in 1993 to develop the National Land Cover Dataset (NLCD). The NLCD 2001 provides accurate standardized non-agricultural ground training data to improve the CDL's representation of the non-agricultural domain.

The availability of commercial off-the-shelf software to perform the land classification and image manipulations has also advanced the remote sensing program at NASS. The addition of See5 decision tree software, ERDAS Image 9.1 remote sensing software and Environmental Systems Research Institute's (ESRI) ArcGIS, and SAS have all played a critical part in enhancing processing capability and efficiency.

Remote sensing currently performs a central role in NASS's statistical program in the estimation of crop area and yields. With today's software, imagery and IT capabilities NASS has transitioned its remote sensing program from a research effort to a production process. Looking to the future NASS expects to see more rapid development of remote sensing applications from research to operational status and hopes to achieve agency benefits with reduced respondent burden, the development of additional spatially rich data and savings from data collections in traditional surveys.

6.0 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.

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.

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 <http://www.nass.usda.gov/research/Cropland/SARS1a.htm> 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.

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 conf http://www.geoinformatics2009.org/

NASS, 2006. Assessment of TM and AWiFS imagery for cropland classification: three case studies.< http://www.pecad.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".