Selection of Multi-Temporal Scenes for the Mississippi Cropland Data Layer, 2004

Fred L. Shore, Ph.D.
Mississippi Department of Agriculture and Commerce
Jackson, MS, USA
fred_shore@nass.usda.gov

Thomas L. Gregory
National Agricultural Statistics Service
Jackson, MS, USA

Rick Mueller
Research and Development Division
National Agricultural Statistics Service
Fairfax, VA, USA

Abstract— **This paper describes the use of the National Agricultural Statistics Services remote sensing methodology to compare three multi-temporal scene-processing scenarios of the Mississippi Delta Region to produce the Mississippi Cropland Data Layer for 2004. The Mississippi Cropland Data Layer Program, initially started in 1999, uses medium resolution satellite imagery, random sample selection of fields, and field data to produce crop acreage estimates and a categorized data layer of the state showing the land use. The focus is on the major field crops and allows a visualization of the entire State of Mississippi land use by year when cloud-free imagery is available. The results for the three multi-temporal strategies show use of 2004 Landsat 5 scenes moderately early (5/6/04) and late (9/27/04) in the season give the superior classification for the Mississippi Delta and the 2004 Cropland Data Layer based on multi-temporal classification statistics and observation of the classified images. On the Mississippi Department of Agriculture and Commerce web site (http://www.mdac.state.ms.us/) you can currently see the changes in Mississippi farming per year from 1999 to 2003. In July 2005, the 2004 data will be added. The Cropland Data Layer results compare favorably with the official NASS Mississippi crop estimates for the major field crops in 1999-2004.**

Keywords-component; Landsat 5; Cropland Data Layer; Mississippi Delta; multi-temporal; crop acreage estimate

I. Introduction

The National Agricultural Statistics Service (NASS) has developed a Cropland Data Layer (CDL) system [1] to classify the location of crop cover types and estimate the acres of each crop [2]. This system has been used and refined for Mississippi since 1999, in partnership with the Mississippi Department of Agriculture and Commerce (MDAC), and Mississippi State University (MSU). State level crop area estimation with satellites began in the early 1990's in the Mississippi Delta [3,4]. These studies showed that the remote sensing regression method indications had less variance than the corresponding direct expansion indication and variance.

Satellite images from Landsat 5 (in prior years Landsat 7 was also used), cloud-free if possible, were selected using the USGS browser viewer (http://glovis.usgs.gov/). These images have 30-meter resolution, 185 kilometers swath width, 7 channels and 16 day repeat coverage. This is a practical resolution allowing field determination on the size of 10 acres or more to be done quickly and is capable of accurately estimating agricultural land use over a state. Alternative choices for imagery such as the Indian Research Satellite (IRS) Advanced Wide Field Sensor (AWiFS) are being evaluated in case of failure of Landsat 5.

The segments for field sampling are approximately 1 square mile in size and are randomly selected by NASS, using an Area Sampling Frame that is based on land use stratification of the state. Fig. 1 is the Mississippi Area Sampling Frame established in 1999. A total of 356 sample segments were selected for the 2004 NASS June Agricultural Survey (JAS). Generally, 20 percent of the segments are rotated in and out of the sample per year. The high (>75%) cultivation stratum in the Delta leads to a concentration of segments in this area allowing for accurate acreage estimation over the Delta region.

The JAS is conducted in the first two weeks of June when all segments are visited by a field enumerator. Segment locator maps are prepared in the Mississippi State Statistical Office (SSO) using ArcGIS. The data layers used for these maps are the segment boundaries generated in the NASS Fairfax office, the geographic information such as the National Forests from www.maris.state.ms.us, and the county map CAD files obtained from the Mississippi Department of Transportation. The major advantages of these maps are the small size (11" x 17"), readable font size street names, and locations of churches and bridges in the immediate vicinity of the segments. In addition to the segment locator maps, the enumerators were provided with 1:8,000 scale aerial photos to aid in locating the field boundaries. The field enumerators record the crop and acres for each field in the segment on the photos and on the field survey questionnaires.

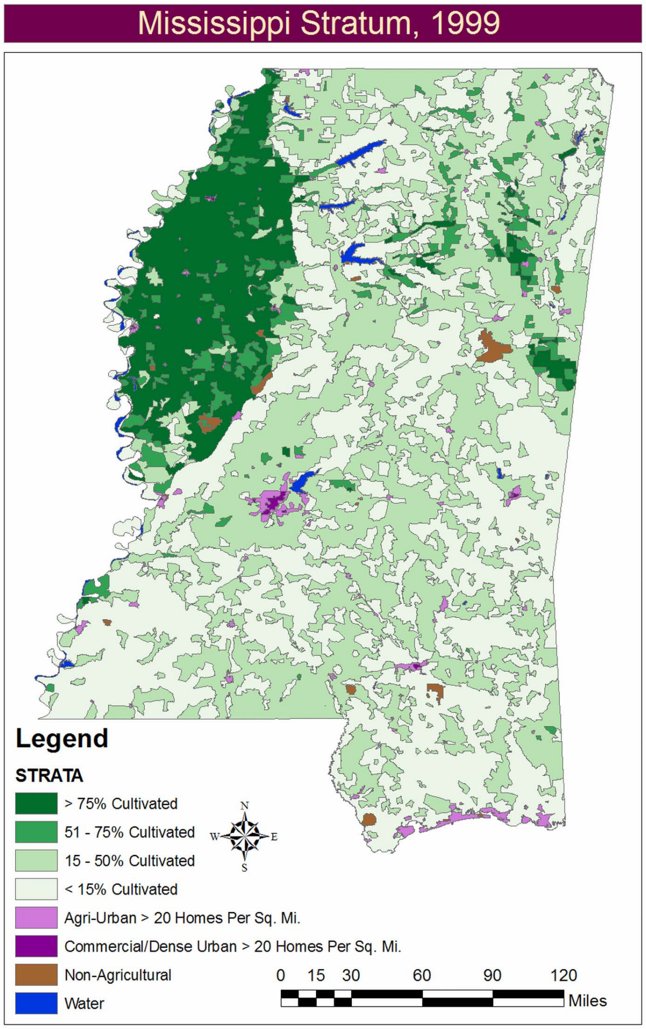

Figure 1. Mississippi Area Sampling Frame used in 2004

The aerial photos are then digitized in the Mississippi SSO using public domain NASS software programs Peditor and the Remote Sensing Project (RSP) data manager [5,6]. Peditor was developed in the 1970's to process satellite imagery to estimate crop acreage over a large area [7], and Peditor's routines are documented [8]. RSP was developed in the 1990's to assist with the digitizing efforts in the NASS State Statistical Offices. RSP allows for heads up digitizing performed on each segment using Landsat imagery as the reference. RSP manages the remote sensing digitizing production workflow for the segments in the JAS.

The optimum Landsat scenes are selected across the state based on image availability, clouds, and planting/emergence of the crops. Two scenes are preferable to accurately separate crops and non-agricultural land use, where a spring and mid-summer scene are considered the optimum for crop identification. The scenes are compiled by analysis district (AD) based on date of image observation, with scene dates the same within an AD, and each AD can be cut by county or scene overlap.

II. The 2004 Procedures

In 2004, Mississippi had an abnormally warm early spring. The weather allowed favorable conditions for most farmers to plant some crops up to a month early. The growers then experienced above normal precipitation until harvest. At harvest, the weather cleared and Mississippi had excellent crop yields including record yields for cotton, rice and corn for grain.

A. Scene Selection

In 2004, NASS was able to get cloud-free Landsat 5 scenes over the Mississippi Delta on 3 separate dates resulting in 3 different multi-temporal processing possibilities. The Landsat 5 scene dates chosen for processing were: 4/4/04, 5/6/04, and 9/27/04, resulting in 3 AD scenarios to choose the best AD for the Mississippi Delta. For the remainder of the State, the central AD had a uni-temporal or single observation date of 9/20/04 and the eastern AD had multi-temporal observation dates of 4/6/04 and 5/8/04. The segments were then clustered by cover type and classified using Peditor's Maximum Likelihood classifier. Once the signatures were cleaned up and the small-scale classification completed, the entire scenes were classified and combined in a mosaic of the entire state. The acreage estimation was performed on the full mosaic of the state. The maps were further processed using ArcGIS software to allow custom map preparation.

B. Data Processing

Peditor and RSP both process data on a modern personal computer. A dual processing computer was used to allow additional speed in scene reformatting and further processing the data.

Processing the 14 bands of each of the scenes in an AD were used to compare pixel intensities and modified ISODATA clustering was performed which allows for merging and splitting of clusters [9]. This was based on field data to give categories that were characteristic signatures for each crop. Each multi-temporal scene gave unique signatures and classifications of the land use. By having two dates there was a greater likelihood of identifying the crop correctly. The software also allows the option of not using every training field in the processing. For example, a farmer may have obviously changed his mind about a crop and planted another or not planted at all, weather may have prevented crop planting, or field boundaries may have changed significantly. Removing the anomalous fields from training was a key component of getting good crop signatures.

TABLE I. STATISTICS COMPARISON FOR 2004 MULTI-TEMPORAL CLASSIFICATION OPTIONS FOR THE MISSISSIPPI DELTA

Analysis District	Dates	Percent Correct (%)				Commission Error (%)		
		Overall	***Cotton***	***Soybeans***	***Rice***	***Cotton***	***Soybeans***	***Rice***
AD01	4/4/2004 and 5/6/2004	81.87	85.78	89.96	94.13	5.08	6.50	27.97
AD06	9/27/2004 and 4/4/2004	80.92	94.61	81.19	96.68	5.79	3.30	22.53
AD08	5/6/2004 and 9/27/2004	85.99	97.19	95.61	90.16	3.23	3.24	13.21

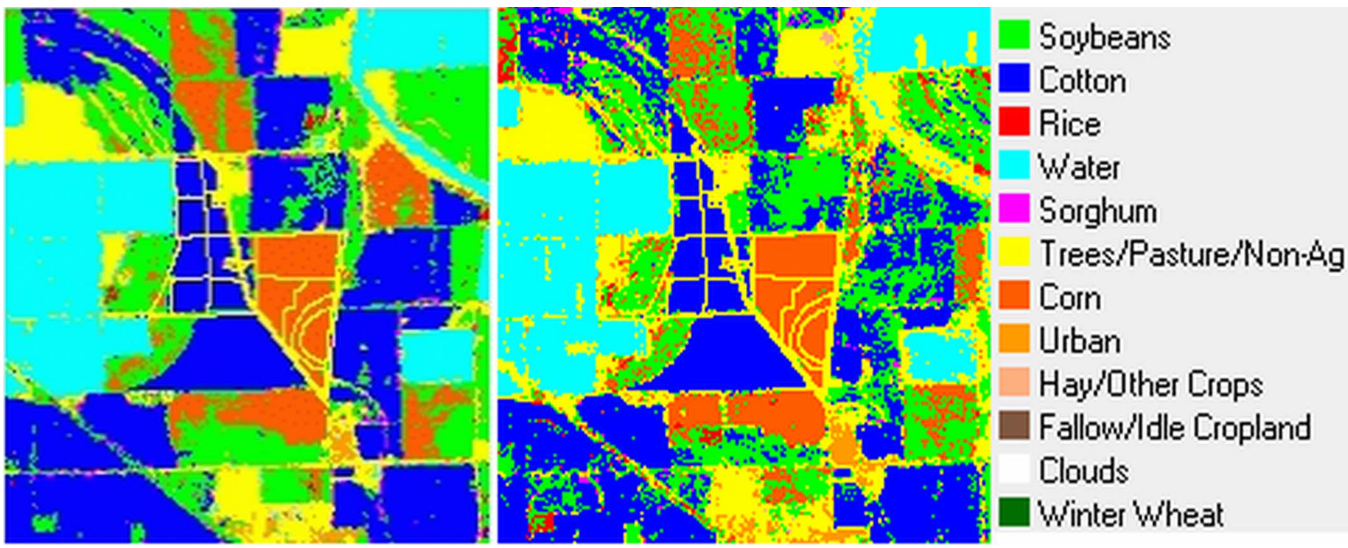

Figure 2. Comparison of AD08 (left) vs. AD06 (right) classification of a parcel of the Mississippi Delta

The 3 multi-temporal AD options for the Mississippi Delta were processed separately and the statistical results for the training data set were obtained for comparison to decide the best option.

Once the final AD scene assignments and mosaic were made, the estimation programs were run. Regression of the assigned pixels was done comparing the training data pixels and grower reported acreage for each AD for the major crops. For areas without sufficient data, such as cloud-covered areas, a supplement from direct expansion using the June Agricultural Survey data was used to obtain a result for the entire state.

III. RESULTS

The improvement in classification using the AD08 multi-temporal analysis district vs. the AD06 analysis district is shown for a portion of Mississippi Delta in Fig. 2.

The classification statistics for the three multi-temporal strategies for the Mississippi Delta are shown in Table 1. Percent correct is the percent of pixels of a given cover type in the training data correctly classified. Commission error is the percent of pixels in the training data classified to the wrong cover type. Use of the lowest commission error to decide the best analysis district shows that AD08 is the best choice for all three crops. Use of the highest percent correct also gives AD08 clearly the best for cotton and soybeans and over 90% correct for rice.

The analysis district AD08, using scenes moderately early and late in the season, thus gives the superior classification for the Mississippi Delta and was used for the 2004 Cropland Data Layer. The usual optimum satellite observation times for cotton and rice estimation in the Mississippi Delta require an early scene taken around the first week of May to the first week of June while the later scene should be acquired around late July to mid-August [10]. In 2004, the spring/early summer scenes were available and the 5/6/04 scenes were best, verifying the historical result. The later scenes were not available until 9/27/04 because of cloud contamination issues. On the Mississippi

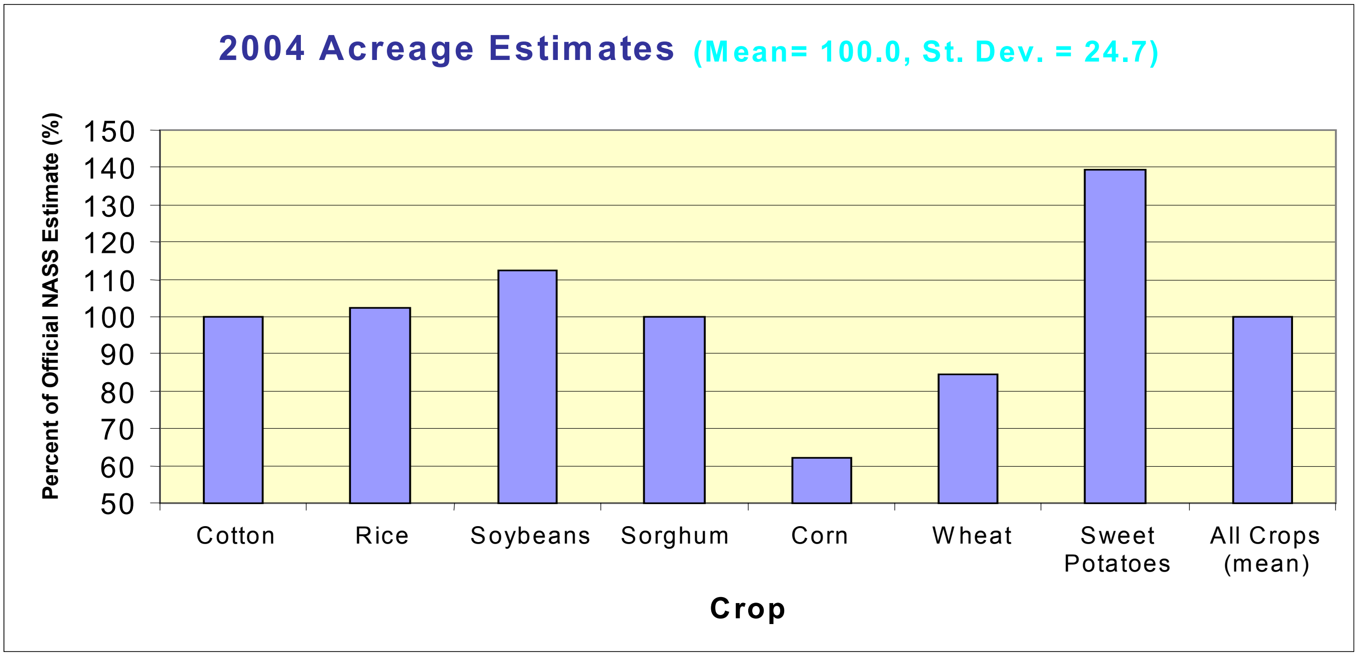

Figure 3. Comparison of CDL acreage estimates with the NASS official estimate

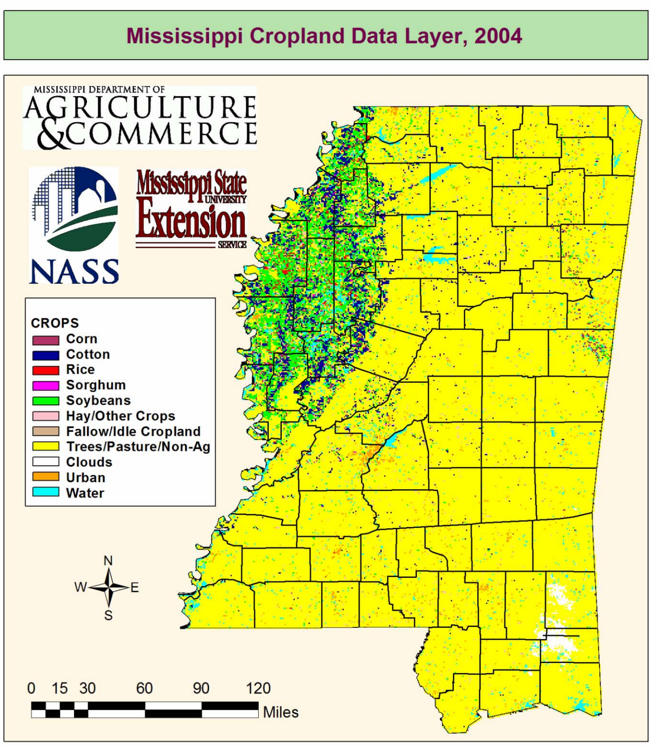

Figure 4. The Mississippi Cropland Data Layer, 2004

Department of Agriculture and Commerce web site (http://www.mdac.state.ms.us/) you can see the changes in land use in the Mississippi Delta per year from 1999 to 2003. The 2004 CDL data will be available in July 2005, after the NASS official county estimates are released.

Fig. 3 shows the results of the CDL as a percentage of the official NASS estimate for the 7 crops estimated in 2004. Generally the major crops and larger acreage crops give the best CDL estimates except the excellent classification results for sorghum shows that even a minor crop can give a good estimate if the planting date and scene selection dates are optimum and enough training data is available.

The results for 2004 for the major crops (cotton, rice, and soybeans) are consistent with the 6-year CDL mean of 100.3% of the official NASS estimate and standard deviation of 7.09.

The Cropland Data Layer estimates were provided to the Mississippi SSO office in December, 2004 to help as an indicator in the preparation of the NASS official estimates for the State of Mississippi, released in early January, 2005.

IV. DISCUSSION

The use of the AD08 multi-temporal analysis district gave the best classification of the crops of the Mississippi Delta and was used for the Mississippi Cropland Data Layer.

The scene selection (availability) is a critical component of the multi-temporal processing required to train the crop classifier. In 2004, the analysis district AD08 had the best dates for scenes as shown by the crop classification statistics

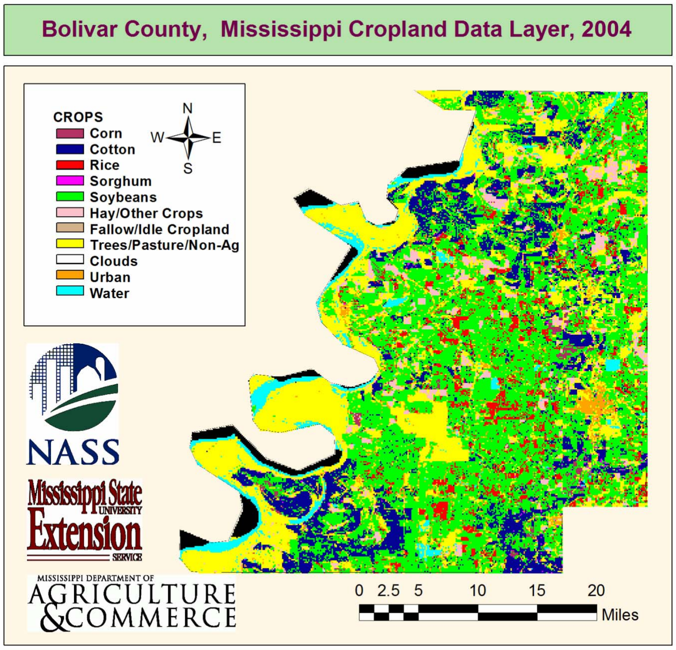

Figure 5. The Bolivar County Cropland Data Layer, 2004

on classification of the training data for the major crops and overall as shown in Table 1.

Soybeans historically give the highest variability of the major crops and for 2004 gave the greatest difference from the official NASS estimate. Differences in planting dates for the soybean varieties and double cropping with winter wheat contribute to the difficulty. With the presence of soybean rust, the use of the soybean CDL for locating soybean land use may become even more important.

For the 2004 data shown in Fig. 3, corn is an example of a crop planted very early and sweet potatoes an example of a crop planted very late. The CDL multi-temporal dates chosen resulted in an under estimation of the corn and over estimation of the sweet potato acres. Wheat is also planted early and was underestimated by the CDL system in 2004.

Fig. 4 shows the 2004 Cropland Data Layer. This map compares favorably with the historical land use stratification map of Fig. 1 and verifies the use of this stratification in the selection of the Area Sampling Frame. The scale of the map for Bolivar County, Fig. 5 shows the classification in more detail.

V. CONCLUSIONS

Acreage estimates using the Cropland Data Layer trend well over time with the NASS official estimates. Three multi-temporal scenarios were tested over the Mississippi Delta for 2004, and AD08 demonstrated the best potential for the major crops of rice, cotton and soybean. The 2004 remote sensing indications performed well when measured against the Agricultural Statistics Board for the major crops. Also, smaller acreage crops can be correctly identified with adequate training data and optimum image acquisition dates, as proven with sorghum in 2004.

The use of the JAS as input to the Cropland Data Layer program provides a synergistic effect across these programs. It not only provides acreage estimates, and an ortho-rectified mosaicked image product, but it reduces the overall estimate bias when the regression estimator is used in conjunction with the JAS.

NASS, MDAC and MSU are committed to the continuation and growth of the Cropland Data Layer program. Despite the lack of adequate satellite coverage in the medium resolution marketplace today, alternative image providers are being investigated to ensure the continuation of this program. The IRS AWiFS satellite images currently provide the most cost effective and temporally sound approach that this program needs to continue operations.

ACKNOWLEDGMENT

Commissioner Lester Spell, Jr., D.V.M., MDAC; Dr. Joseph H. McGilberry, Director, Mississippi Cooperative Extension Service; James Brown, Mississippi Department of Transportation; and the NASS Field Enumerators in Mississippi were critical to the success of this project.

REFERENCES

[1] R. Mueller and M. Ozga, "Creating a Cropland Data Layer for an entire state," Proceedings of the ACSM-ASPRS 2002 Conference, Washington DC, ASPRS, Bethesda, MD, April 2002.

[2] G. Hanuschak, R. Hale, M. Craig, R. Mueller, and G. Hart, "The new economics of remote sensing for agricultural statistics in the United States," Caesar Conference, Rome, Italy, June 2001.

[3] M. E. Craig, "The NASS Delta Project: 1991-92 Rice and Cotton Acreage," handouts from presentation at 1993 ACSM/ASPRS Annual Convention, New Orleans, Louisiana, February 1993.

[4] M. L. Graham, "State level crop area estimation using satellite data in a regression estimator," Proceedings of the International Conference on Establishment Surveys, ICES I, American Statistical Association, Alexandria, Virginia, pp. 802-806, June 1993.

[5] J. D. Allen, "A look at the remote sensing applications program of the National Agricultural Statistics Service," Journal of Official Statistics, vol. 6, pp. 393-409, 1990.

[6] J. D. Allen and G. A. Hanuschak, The Remote Sensing Applications Program of the National Agricultural Statistics Service: 1980-1987, U.S. Department of Agriculture, NASS Staff Report No. SRB-88-08, 1988.

[7] M. Ozga, M. E Craig, "PEDITOR - Statistical image analysis for agriculture," Washington Statistical Society (WSS) Seminar, April, 1995.

[8] C. D. Day, A Compilation of PEDITOR Estimation Formulas. U.S. Department of Agriculture, NASS/RDD Research Paper No. RDD-02-03, January 2002.

[9] M. Bellow and M. Ozga, "Evaluation of clustering techniques for crop area estimation using remotely sensed data," Proceedings of the American Statistical Association, Survey Research Methods Section, Atlanta, Georgia, pp. 466-471, 1991.

[10] J. D. Allen, Remote Sensor Comparison for Crop Area Estimation Using Multitemporal Data, U.S. Department of Agriculture, NASS/R&AD SRB Staff Report No. SRB-90-03, 1990.