The New Economics of Remote Sensing for Agricultural Statistics in the United States

by
George Hanuschak, Robert Hale, Michael Craig and Rick Mueller
National Agricultural Statistics Service, U.S. Department of Agriculture
and
Galen Hart
President, Galen Hart Associates**,** Virginia, USA

Keywords: remote sensing, agricultural statistics, cost, economics, partnerships

Executive Summary

The National Agricultural Statistics Service started an applied research program in the early 1970's to utilize space-borne satellite imagery and digital data, such as Landsat, to identify crop type and measure crop acreage for selected States in the U.S. In the initial research, software development for information extraction was the key activity. NASS needed the traditional remote sensing outputs of cover type and probability of being correct. In addition, NASS needed an unbiased (or a negligibly biased) statistical estimator of crop area at the State and county level. Using Purdue University=s LARSYS (early 1970's version) system as the base, NASS and the University of Illinois=s Center for Advanced Computation developed a customized software package called EDITOR to accomplish these tasks. NASS staff converted the software across several platforms and called it PEDITOR. PEDITOR gives conventional remote sensing categorization (or classification) outputs and a double sampling regression estimator as well. Over the years, NASS staff have continually improved PEDITOR by adding functionality and efficiencies.

Most recently, NASS staff have emphasized system efficiency with expert system like features to ease the labor burden of analysts. In addition, recent improvements in the county level estimates have been added as well. The system now runs on high end Windows NT desktop computers. The program is currently at seven States plus a minority student outreach program with Florida A&M University. In the last several years, NASS has developed a Cropland Data Layer (CDL) in geographic information system (GIS) format (See Figure 1) for public use. It is being used by GIS proficient users for watershed monitoring, agribusiness planning, prairie water pothole monitoring, crop rotation pattern analysis and animal habitat monitoring. The CDL can be viewed and ordered on CD at http://www.nass.usda.gov/research/Cropland/SARS1a.htm

The Old Economics (1970's-1980's)

The economics has changed rather substantially over the last three decades. The initial statistics and software research in the early 1970's was dominated by research staff costs and computer processing costs. The first full State project, Illinois in 1975, cost $750,000 but still included

substantial research and statistical and software development costs. The first applications costs were established for processing the State of Iowa (1978) in time for end of season crop area estimation. At that time, the total cost per State was estimated at $300,000 and the data source was Landsat Multi-Spectral Scanner (MSS) at 60-80 meter resolution, with 4 spectral bands of data.

By the early 1980's, there were eight States being processed. The total cost per State was $150,000. Major item costs were associated with remote sensing analyst staff (although spread out over 8 States which increased efficiency substantially over one or two States with about the same staffing levels) and the cost of processing the data on mini-computers and on the ILLIAC and Cray supercomputers. Data costs were $300 per Landsat scene and generally a small portion of the total project cost. When Landsat thematic mapper became available in 1982 at 30 meter resolution with 7 spectral bands of data, the program faced a fork (applications continued or a research program and switch to Landsat TM) in the road and only had resources to go in one chosen direction. At first, NASS continued the eight State application using Landsat MSS data, with a small amount of resources devoted to research on Landsat TM. Even with the limited research program resources, it became apparent that the accuracy levels could be improved substantially by using Landsat TM but at a higher cost for the data and for the information extraction (more bands and better resolution). Thus in 1988 due to budget cuts and U.S. government plans to not have MSS data on future Landsat or other earth resource observing satellites, NASS abandoned the eight State application using Landsat MSS but increased the research program with Landsat TM.

A new applications program was initiated in 1991 in Arkansas and Mississippi. In 1992 Louisiana was added to the program. Budget cuts hit again and Louisiana and Mississippi were dropped from the program. In fact in 1995, the program was only intact for Craighead county in Arkansas. Needless to say is that the program had a very rocky road at that point in time. Thus, it was apparent that NASS had to partner and resource share with other Federal and State government Agencies in order to continue and expand the program again and add a public cropland data layer for distribution as well as the internal statistical products. In preparation for this, NASS analyst staff emphasized a more automated and more expert system like version of PEDITOR to ease the labor burden on analysts. When this was adequately accomplished, NASS was in position to seek out Federal, State and University partnerships.

Resource Leveraging Partnerships

Starting in 1997, a new and key data license partnership between NASS and USDA=s Foreign Agricultural Service (FAS) and the Farm Services Agency (FSA) was established. Landsat 5 data could be used for projects that serviced these agencies programs. Thus, the Landsat data costs for a program expansion were reduced. In addition, with Landsat 7, data costs at EROS Data Center were set at $600 for a system geo-referenced data set without licensing restrictions. USDA FAS has a USDA-wide license for Landsat 7 with Radarsat for $405 per scene. Thus, the Landsat data costs were no longer a major constraint as when the cost for Landsat 4&5 peaked for government buyers at $4,400 per Landsat scene in the early 1990's.

The highest cost historically was for expert remote sensing and statistical analysts to run the

complex PEDITOR package for both remote sensing outputs and for statistical estimates with relative sampling error outputs as well. NASS had a small group of expert analysts in its♠ Research Division who were doing centralized analysis for several States. However, it was recognized that a decentralized analysis staff directly in NASS State Statistical offices would expand the Agency♠s analytical capabilities. In addition the analyst will have the advantage of more localized knowledge of the crops and cropping practices and other sources of data to evaluate the relative contribution of the Landsat to the NASS crop area estimation program at the State and county level.

A local analyst (with a stable position and longevity) and State office management will be in a better position to service other State government and university and farm organizations and agribusinesses by providing them a Cropland Data layer in geographic information system (GIS) format. For example, the analyst in Illinois did both Illinois and Indiana in a regional concept which is the most cost efficient. Small two to four State regions will likely be the most cost efficient model.

This data layer has already proven valuable to data users of NASS who often combine the NASS cropland data layer with other data layers. Some examples are for water quality assessments for watersheds, and location plans for a new agribusiness facility, such as a soybean crushing plant and for grain storage and transportation planning and for prairie water pothole monitoring in North Dakota**.**

Since 1998, NASS has added the States of Illinois, Indiana, Mississippi, New Mexico, and Iowa and continued to do North Dakota and Arkansas as well. Florida A&M University and NASS also entered into a cooperative program for minority student outreach in remote sensing and GIS by analyzing a portion of northern Florida. Establishing and maintaining effective and win-win partnerships for the Cropland Data Layer is challenging though. The analyst position needs to be strongly supported by all the partnering organizations and needs some longevity (5 years or so) to be effective. Some of the partnerships are stronger and more stable in terms of longevity than others. A companion paper by Craig (ASPRS 2001) entitled ■ A Resource Sharing Approach to Crop Identification and Estimation • goes into more detail about each partnership for the Cropland Data Layer program.

The New Economics (late 1990's and 2000)

NASS has entered into a new economics for crop area estimation and a Cropland Data Layer for public consumption through the new resource leveraging partnerships. The total cost per State is now $75,000 and should drop further with the addition of more States. See the attached Figures 2 and 3 for the graphic representation of the cost reduction over time (inflation adjusted and non-inflation adjusted). Major reductions were realized in the cost of the Landsat data and even more dramatic in the cost of data processing which has gone from mini-computers and supercomputers to high end PC desktops now. All the processes, including full Landsat scene multi-temporal pixel categorization, are done on PC♠s The non-inflation adjusted total project cost has been cut in half (See Figure 2). When adjusting for inflation (See Figure 3), the cost reduction is more than fivefold from the first large scale application in the early 1980's. The current program and

the early 1980's are the only comparable periods to compare costs over. The reason is that the number of States that labor costs are being spread over are very similar in number and scope of effort. Dramatic cost reductions have been achieved in the cost of the Landsat data and for the data processing of the Landsat data to categorize it into most likely crop types (See Figure 4). Spreading out the labor costs over as many States as possible and having co-funding partners is the key to future cost reductions.

Perhaps more importantly, there are now more accurate and valuable output products, both internal to NASS and external to NASS data users. These include a State level estimate with relative sampling error reductions on the order of two to five compared to the ground sample data only. County level (small area) estimates with measurable sampling errors which have been recently improved with several alternative estimators. Third, a new publicly available Cropland Data Layer at the 30 meter pixel level categorized by crop type and formatted for standard geographic information systems input.

The use of the Cropland Data Layer by GIS proficient data users outside of NASS is a major new thrust. In North Dakota, the NASS State Statistical Office (SSO), the North Dakota State University and an agri-business firm used the Cropland Data Layer as one input into a key agri-business decision. The location of a soybean processing plant was being evaluated. By combining the Cropland Data Layer with a transportation network layer and other layers such as commercial grain storage facilities for soybeans, a decision model was built by the University and used in the agri-business decision for the plant location. Another anticipated major use of the Cropland Data Layer is in watershed monitoring and in fact is one of the major reasons for the North Dakota State government interest in the program. The State of Illinois chose to add to the Cropland Data Layer by adding other ground data on land covers and converting it to a Land Cover Layer.

Private Sector Perspective

The economics of private sector remote sensing for agriculture and renewable natural resource monitoring has followed a somewhat similar pattern to the USDA/NASS experience with a peak in the 1980's followed by a downturn and then a substantial recovery at the turn of the century. In the late 1970's and early 1980's, a number of commercial applications were developed. Many of these applications dealt with clustering and classification of Landsat Multi-Spectral Scanner (MSS) data (60 meter to 80 meter resolution) into land cover types, including agricultural land covers in some cases.Many of these applications produced large scale land cover maps for government and private sector use. However, a number of these products had limited accuracy assessments and were often restricted by the lack of a statistical sample of ground data for verification purposes. Usually, a small amount of ground data or photo-interpreted aerial photography was used for verification and creation of the classification matrix (omission and commision errors).The attempted commercialization of Landsat, in the early 1990's, came next and sharply increased prices per Landsat scene (over $4,000). A number of the firms with Landsat applications such as large scale land cover were forced to scale back or, in some cases, go out of business. Several firms diversified further into value added products that went beyond classification and mapping. Some were as simple as enhanced raw or classified image products put on glossy color prints. Other value added products were more complex, such as using

geographic information systems to combine Landsat data with other data layers to customize products/solutions for clients.

Among the better known remote sensing, geospatial information firms in the Unites States are Earth Resource Satellite Corporation and Pacific Meridian (now a unit of Space Imaging Corporation). These geospatial information companies, along with many others now, often ▪mix• multiple data sources in GIS format to customize products/solutions for their clients. For example, one could combine weather data contours (precipitation, temperature, wind etc.), soils data contours, satellite vegetative index data and crop condition data to get a crop yield forecast. Some of the data sources are satellite image data, aerial scanner digital data, U.S. Geological Survey digital maps, digital terrain maps, digital soils maps, official government statistics and ground data with global positioning systems (GPS) locations etc.

With the launch of Landsat 7 in 1998, a new era began with government subsidized data pricing at $600 per scene, system registered and license free. This was a big boost to all Landsat applications, both public sector, private sector and universities benefitted from the reduced prices. Large scale applications took off again as large area studies were no longer severely restricted by data prices. At the recent Landsat Data Continuity Mission Workshop, held at the U.S. Geological Survey and co-sponsored by the National Aeronautics and Space Administration (NASA), a number of data distributors,users and value added firms reported new growth opportunities created by Landsat 7. One specific value added firm example presented at the conference stood out as Metapath Software International (MCI) reported an annual revenue increase of 277% over the previous year. MCI is a value added firm that does worldwide urban analysis with Landsat 7 (pan-sharpened and license free) as one of the major ingredients in there geospatial information extraction process. Many of those attending the Workshop expressed a desire for Landsat 8 to mimic the total success of Landsat 7, that is high quality data, reasonable pricing, similar spectral coverage, near nadir views , fixed polar orbit and wide swath coverage.

Agribusiness, particularly the large international commodity firms, are using remote sensing as one input for crop forecasts and estimates in statistically ▪under served• production areas of the world. Their interests range from their own forecasts of large area produced crops (grains, fiber and oilseeds) to more limited area crops (palm oil, cacao, tea and coffee). Their expenditures for this type of information development are largely unknown due to the proprietary nature of their business. However, those who have experience in the area agree that hundreds of thousands or more dollars are spent annually by these firms for their geospatial information extraction, including crop forecasting and estimation.

Similar to the described USDA/NASS experience, more frequent temporal coverage would benefit private sector geospatial information extractions firms with agricultural clients. Data coverage is a major issue in timely crop monitoring applications. In many rain fed crop areas of the world, better than 16 day coverage is required. In some tropical areas, radar data is being explored because of the lack of optical sensor data coverage due to excessive cloud cover during fixed satellite overpasses. Three potential future data providers (Resource 21, Matra Marconi and RAPIDEYE) have identified agriculture as a primary commercial market. If any of these ventures succeed and have reasonable pricing structures, then that should be another boost for agricultural applications of space-borne remotely sensed data.

Satellite Data Sources

NASS staff have a very strong preference for the Landsat data. The nadir look, the wide area image swaths, and the spatial resolution and spectral bands are virtually ideal for the NASS Cropland Data Layer and crop estimation program. The only issue with Landsat is the temporal frequency. NASS staff desires eight day coverage as a minimum and would prefer even better temporal coverage. Backup to the Landsat for the NASS program is the Indian IRS satellite series, which have many Landsat-like features. SPOT has not proven to be a cost effective backup for the NASS program. NASS has used SPOT for area sampling frame construction, on occasion, where price reductions are available through State consortiums or now under a USDA-wide license or existing credits from USDA/FAS.

NASS gets questions about its♠ plans for using very high resolution satellite data. At this time there are no major plans to use very high resolution data, such as IKONOS in the NASS program other than for occasional small quality control checks. NASS has farmer reported data from a stratified area frame sample to evaluate the Landsat data. The swath width of IKONOS is too small for wall-to-wall coverage needed for the Cropland Data Layer and the cost per square mile too high relative to Landsat. This only covers NASS needs as the IKONOS satellite provides very high quality data for many site specific applications, such as mapping a city.

Future Perspective (2001-2010)

NASS will continue to pursue partnerships, primarily with State governments, to expand the crop area estimation and Cropland Data Layer program to more States. The desires are for program expansion to the top 15-20 total cropland States in the United States. The Midwest , the Delta and the Great Plains are the likely areas for expansion as resources allow for. Expansion beyond that point is unlikely.

The value of the Cropland Data Layer to the general public is hard to quantify but is considered quite substantial to those proficient GIS data users who combine it with other data layers to solve their problems of interest. Examples of uses to date are watershed monitoring, prairie water pothole monitoring in the Dakotas, grain transportation and storage planning, animal habitat monitoring, agri-business plant location analysis, farm equipment dealer planning and crop rotation pattern analysis.

Acknowledgements

The authors would like to extend their thanks and appreciation to all the analysts in the States involved. Those are Patrick Willis in Illinois, Dath Mita in North Dakota, Ken Akins in Mississippi, and Dave Christiansen in New Mexico. In addition, the State Statisticians and their staffs for extensive ground data preparations for the States in the program. Thanks also go to Martin Ozga for recent improvements in PEDITOR aimed at reducing analyst labor, including a new mosaicking algorithm. Thanks to Jim Mergerson, Charles Day and several members of the Area Frame Section for assisting in analysis preparations or analysis.

The authors would also like to extend their appreciation to Carol House, Fred Vogel, Joe Reilly, Rich Allen and Ron Bosecker of NASS for their headquarters executive management support of this program and to Allen Vandergriff of USDA/FAS and Glenn Bethel of USDA/FSA.

References

Allen, J. D., and Hanuschak, G. (1988). The Remote Sensing Applications Program of the National Agricultural Statistics Service: 1980-1987. SRB Staff Report Number SRB-88-08 USDA, NASS, Washington, DC-USA

Bellow, M. (1994b). Large Domain Satellite Based Estimators of Crop Planted Area. In: *1994 Proceedings of the Section on Survey Research Methods*, August, 1994, American Statistical Association, Toronto, Canada, pp.194-199.

Craig, M. (2001). A Resource Sharing Approach to Crop Identification and Estimation. *2001 ASPRS/ACSM Annual Convention Technical Papers, available on the CD, Proceedings of the ASPRS 2001 Conference*, Bethesda, MD-USA.

Day, C. (2001). A Compilation of PEDITOR Estimation Formulas. RDD Research Paper RDD-01-01, USDA, National Agricultural Statistics Service, Washington, DC-USA.

Fishman, J. (2001). The Landsat 7 Applications of Metapath Software International (MCI), USGS/NASA Landsat Data Continuity Workshop, Reston VA-USA.

Graham, M. (1993). State Level Crop Area Estimation Using Satellite Data in a Regression Estimator. In: *Survey Methods for Businesses, Farms, and Institutions*, ICES Part I, NASS Research Report No. SRB-93-10, USDA, Washington, DC-USA, pp. 7-11.

Hanuschak, G., Allen, R., and Wigton, W. (1982). Integration of Landsat Data Into The Crop Estimation Program of USDA's Statistical Reporting Service. *Proceedingsof the 1982 Machine Processing of Remotely Sensed Data Symposium*, Eighth International Symposium , LARS, Purdue University, West Lafayette, IN-USA, pp. 45-56..

Mueller, R. (2000). Categorized Mosaicked Imagery from the National Agricultural Statistics Service Crop Acreage Estimation Program. Available on the CD: *Proceedings of the ASPRS 2000 Conference*, ASPRS, Bethesda, MD-USA.

Ozga, M., and Craig, M. (1995). PEDITOR - Statistical Image Analysis for Agriculture. Presentation at the Washington Statistical Society (WSS) Seminar, April, 1995, USDA, National Agricultural Statistics Service, Washington, DC-USA.

Ozga, M. (2000). Batch Processing of Remote Sensing Jobs on the PC. Available on the CD: *Proceedings of the ASPRS 2000 Conference,* ASPRS, Bethesda, MD-USA

Vogel, F., and Bange, G. (1999). Understanding USDA Crop Forecasts. In: NASS/WAOB Miscellaneous Publication No. 1554, U.S. Department of Agriculture, Washington DC-USA.

Figure 1: *Cropland Data Layer*
McLean County, Illinois

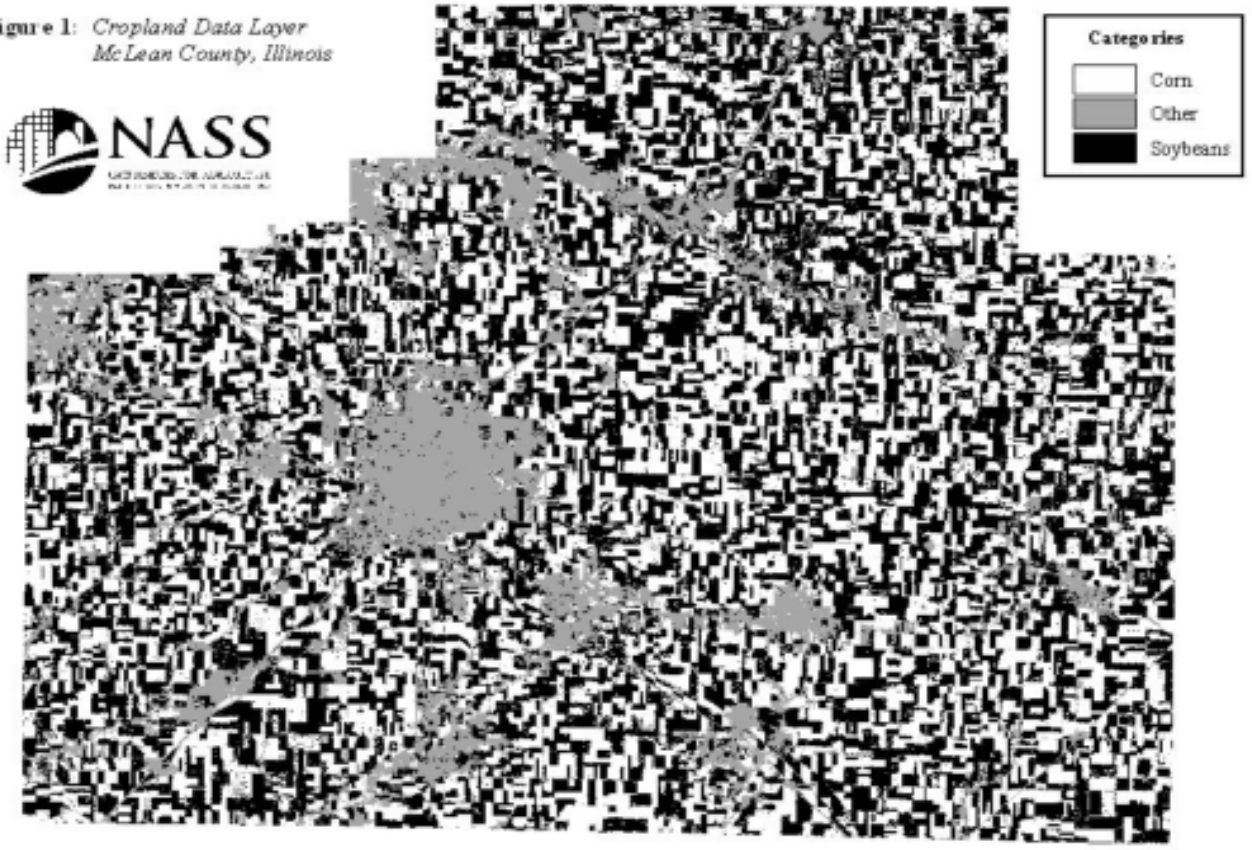


Figure 2: *Cost Per State (Unadjusted)*

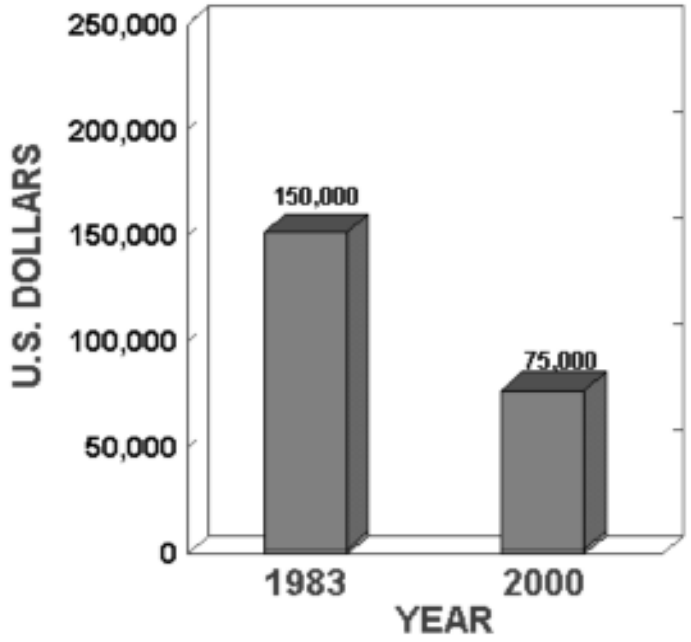
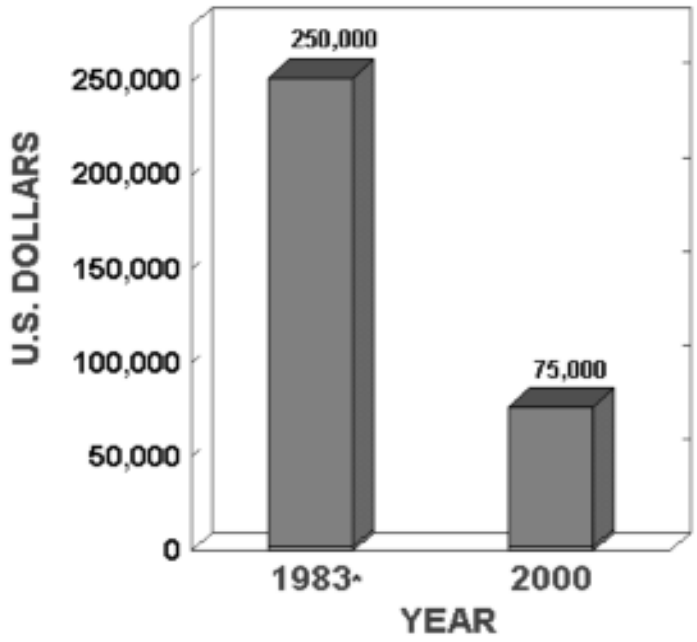


Figure 3: *Cost Per State (Adjusted for Inflation)*



*1983 adjusted to year 2000 dollars

Figure 4: *Major Item Costs (Adjusted for Inflation)*