FAO Regional Project
Improvement of Agricultural Statistics in asia and Pacific Countries
(GCP/RAS/171/JPN)

First Focal Point Meeting of the Project; 7-8 June 1999
Seminar on Remote Sensing for Agricultural Statistics: 9-11 June 1999
Bangkok, Thailand

Session 3, Seminar on Remote Sensing for Agricultural Statistics

Appropriate Role of Remote Sensing in U.S. Agricultural Statistics

By **Mr. Robert Hale**

Head, Area Sampling Frame Section, Research Division, National Agricultural Statistics Service, Department of Agriculture, U.S.A.

THE APPROPRIATE ROLE OF REMOTE SENSING IN U.S. AGRICULTURAL STATISTICS

BY

ROBERT C. HALE
GEORGE A. HANUSCHAK
MICHAEL E. CRAIG

NATIONAL AGRICULTURAL STATISTICS SERVICE
U.S. DEPARTMENT OF AGRICULTURE

I. SUMMARY

The National Agricultural Statistics Service of the United States Department of Agriculture has been utilizing digital earth resource observation satellite data since the launch of the first Landsat in 1972. There are three major applications in the U.S. agricultural statistics program. These are area sampling frame construction, crop area estimation, and crop condition assessment. In addition, we are in the early stages of research on using remote sensing in yield models.

The area sampling frame construction process is the largest operational application. Area sampling frames remain as the statistical foundation of many U.S. probability based agricultural surveys, since area frames have complete frame coverage and no duplication of land areas. NASS has been utilizing area frame sampling since 1954 and nationwide since 1965. The land area of the country (state by state in the U.S.) is divided into broad land use or land cover strata. From 1954 to 1978 aerial photography was the primary source of information for this stratification. From 1978 to the present, Landsat image products have been the main source of stratification information. In 1988, NASS started a joint research effort with the National Aeronautics and Space Administration (NASA) to convert the stratification process from a basically manual image interpretation process to an almost entirely digital process by combining Landsat Thematic Mapper digital data with digital line graph map data from the U.S. Geological Survey. French SPOT data is also being used as a source for delineation of city and urban boundaries. Benefits of the new method include both improved cost efficiency and timeliness.

The second application is the use of categorized Landsat Thematic Mapper data in combination with area sample frame based ground-gathered data to improve the precision of major crop acreage estimates in several states in the U.S. These estimates are calculated in an operational time frame and provided to the Agency's Agricultural Statistics Board as input to the official estimates released by the Agency during the crop season. The well documented regression estimator approach is used. Categorizations of multi-temporal Landsat data are used to calculate regional, state, and county level estimates. In addition, crop specific digital data layers in georegistered format are

provided to the state offices on CD-ROM. The program is expanding through partnerships of both federal and state governments.

The third area of utilization of remotely sensed data involves the use of vegetative indices calculated from National Oceanic and Atmospheric Administration's Advanced Very High Resolution Radiometer (AVHRR) sensor. The low resolution AVHRR data is used to monitor changing vegetation condition over large areas throughout the growing season. Indices from a biweekly composite (of the daily data) are used to measure vegetation vigor; ratio comparisons to the previous year and to a median of several previous years plus within-year thumbnail time series and frost danger products are produced. All of these map products are loaded to the Agency's Internet web page for use by anyone.

Overall, NASS is a fairly extensive user of space based remotely sensed data and related spinoff technologies such as GIS in its U.S. Agricultural Statistics Program. However, in relation to the overall NASS mission of providing agricultural statistics on hundreds of items throughout a year, the portion of NASS's program that utilizes remotely sensed data is not large. Especially if one compares it to the sometimes seemingly "utopian and panacea type" claims of many in the Remote Sensing Community in the 1970's, then it seems like a fairly modest (albeit an important one) addition to the U.S. Agricultural Statistics Program. The Agency has, however, been able to quite successfully integrate and supplement its existing probability based (area, list, and multiple frame sampling) estimation program by utilizing digital and image space based remotely sensed data for the three applications described above.

II. AREA SAMPLING FRAME CONSTRUCTION

The most recent thorough description of the conventional paper based area sampling frame construction process was given by Cotter and Nealon (August 1987 - reprinted June 1989). Earlier thorough descriptions of the paper based process were given by Houseman (November 1975) and Huddleston (April 1976). The first description of the use of Landsat images in the conventional paper based system was given by Hanuschak and Morrisey (October 1977).

More current descriptions of the new system called the Computer Assisted Stratification and Sampling (CASS) system were given by Cotter and Mazur (February 1992) and in two recent FAO reports on Multiple Frame Agricultural Surveys Volume 1 (1996) and Chapter 4 Volume 2 (1998).

NASS has been utilizing area frame sampling in some form since 1954. In order to select an area sample, there first must be an area frame. There are two main inputs needed to construct an area frame as used in the U.S.: some image of the surface features of the land, and a map base of known scale for measurements. For many years, NASS used black and white aerial photography for the image and paper maps. These served us well and can still be used to construct an excellent area frame. If you consider aerial photography as remotely sensed information, we have used remote sensing for nearly 50 years. In 1978, we began using printed satellite (Landsat III) imagery to supplement the aerial photography (the advantage being that the satellite imagery was more current and there was information to be gained from the colors of the imagery).

In 1990, we began the move to a totally digital process for area frame construction. The satellite imagery is displayed on a computer monitor instead of being printed. We have the option of choosing what bands to display, the colors to use, and the scale to view the image. The paper maps have been replaced with digital line graph data or scanned maps. We no longer need to draw boundaries on photos, transfer the boundaries to maps, and then measure the frame. All is accomplished with the click of a mouse. The new technology and methods have lead to better area frames that cost less and are produced with the most up to date information.

The medium resolution (30 m. pixel) of the Landsat is well suited for area frame construction in the U.S. There is a good balance between the detail in the data and the volume of data required to cover a large area. In addition, the spectral bands recorded by the Landsat satellites are well suited for agricultural applications. The higher resolution of, for example, the French SPOT panchromatic imagery can be useful for smaller areas where more detail is required. The use of newer generation satellite data with many spectral bands and very high resolution can sound appealing, but you must consider the huge amount of data that needs to be obtained and processed to cover the area of a country. Acquiring cloud-free high resolution imagery of an entire state or large region is a daunting task.

Using remote sensing data to help construct an area frame certainly has advantages. Generally, the data is very current and most satellite data is georeferenced to fit well into many of the commercially available geographic information system (GIS) software packages. However, it is very difficult for remote sensing data alone to replace the multipurpose role of an area frame. The area frame can be used to collect information on crop area planted and harvested, but it can also be used to collect information on other agriculture items of interest such as livestock numbers, economic variables, and farm demographics.

The area frame in NASS serves three main purposes: 1) to collect planted area data for major crop varieties, 2) to supplement list information in livestock and economic surveys, and 3) to supply ground truth data for our remote sensing crop acreage estimation program. The data from our area frame for the major crops of corn, soybeans, wheat, and cotton can be used alone for accurate U.S. level estimates. With the increased specialization of livestock operations in the U.S., the area frame is used with our list in multiple frame surveys for livestock. A more detailed description of our remote sensing acreage estimation program is next.

III. CROP ACREAGE ESTIMATION

Digital categorization of current year Landsat TM data can also be used to estimate crop acreage. The data collected by enumerators during an area frame survey is an excellent source of ground "truth" data for training the computer classifier. Even so, there is always a certain amount of error inherent in any computer classification. A regression based estimator for crop acreage estimation was developed in the early 1970's that utilizes both the area frame sample segment farmer reported data and the categorized Landsat data. Since purchasing and processing Landsat TM data can be expensive, NASS is only able to use this procedure for selected states. However, this method can add substantial information on crop acreage and related land covers in important crop states or regions.

The types of additional information are indeed impressive, as the confidence band associated with the acreage estimators at the state level can often be reduced by a factor of four or more. In general, these results have been achieved using multi-temporal two date coverage for each scene area of interest, one scene before crop emergence and one scene during peak vegetative state if available. However, due to cloud cover problems and limited number of satellite overpasses, the coverage of the satellite data varies from year to year. The results are not available until the final estimation period after the crop season, due to acquisition and processing time required. The regression estimator does adjust for the varying satellite coverage from year to year by utilizing the area frame farmer reported data alone to estimate for cloud covered area. So even though the areas covered by Landsat may vary from year to year, combining the area frame and the regression estimators consistently covers the entire area of interest such as the entire state or major crop producing areas of a state and significantly reduces the variance of the estimate.

In addition, the crop specific categorization of the Landsat data (30 sq. meters) is available for small area estimation and mapping. Our recent experience indicates that the crop specific categorization, which would be released to the public at large as a digital GIS data layer, is the product of most interest to data users and partners outside the Agency. The tighter confidence band, associated with the regression estimator, is viewed as an internal gain primarily.

The crop specific categorization can be released and be of value to many external data users several months after the crop season, unlike the acreage estimates which must be made during the season or at least by the end of the crop season. Prospective data users for the crop specific categorization include policy makers at all levels of government, farm organizations and farmers, grain transportation and storage industries, value added remote sensing industries, environmental monitoring groups, crop disease and pest monitoring industries, and hydrology monitoring groups, such as the following the growth of prairie potholes in the Northern Plains in the U.S.

In recent years, partnerships are enabling a modest expansion of the program to additional states. In 1997, NASS and the Foreign Agricultural Service/Farm Service Agencies of the USDA entered into a new agreement to share the full resolution U.S. Landsat coverage for internal USDA crop monitoring programs. In addition, in 1999, several state governments and other Federal-State agencies are partnering with NASS to expand the crop specific categorization program. In 1999, Illinois, Arkansas, Mississippi, North Dakota, and New Mexico are anticipated to be part of the program. The partnerships enable cost and staff sharing to accomplish goals that no one of the organizations would likely be funded to do individually. The successful launch of Landsat 7 in April 1999 was a welcome event to this project and undoubtedly to many others around the globe.

IV. VEGETATIVE INDEX MAPPING OVER LARGE AREAS

NASS and Agricultural Research Service (ARS) conducted joint research in the late 1980's that led to the development of large area use of NOAA's AVHRR data from polar orbiting weather satellites for crop monitoring purposes. The program was modeled after the Statistics Canada system that went operational in 1988. NASS was slow to get into the use of coarse resolution remotely sensed data such as AVHRR because it already had extensive data sets on crop stage and condition, farmer reported yields, and objective yield crop counts and measurements. It was doubtful that remotely sensed data could compete with those systems. Those systems have long established time series of data and perform quite well in crop monitoring and estimation.

However, due to the rather complete spatial coverage of the AVHRR data and the daily overpasses, the research staff chose to examine the vegetative indices from the AVHRR data as a supplement to the existing crop monitoring systems. The research staff chose to use a biweekly composite of the Normalized Difference Vegetative Index (NDVI) produced operationally by the EROS Data Center (EDC) of USGS, rather than try to establish a processing center to deal with the daily data. NASS has procured and stored the biweekly composites from 1988 to current. NASS puts the images on its WEB site at http://www.usda.gov/nass/. The images are especially useful for large area and timely views of vegetative condition, when crops are under stress due to lack of precipitation, certain diseases that attack the plant chlorophyll, and large area flooding. The first major event that we observed on the images was the large area flooding in the Midwest U.S. in 1993. The images clearly indicated early in July that large areas of planted corn and soybeans were flooded or saturated with water. Indeed the flooding plus an early fall freeze led to very low final yields for corn and soybeans in the Midwest U.S. that year. Some other major events observed on the NDVI images were a very late planting season in the corn belt in 1996, a major drought of the 1995/96 winter wheat crop in the southern plains, and a large area drought during the summer months in Texas in 1998. NASS has also overlayed contours associated with average dates of first freeze in the fall to see what amount of area might still be vulnerable to freeze damage. USDA policy makers were quite impressed with being able to view these rather important events on a one page graphics display.

NASS is also careful to point out the limitations of the data as well as the strengths. For example, the spatial resolution of one square kilometer limits the data to non crop specific applications. That is why large area events are the best observed by this type of data. A large area drought often effects all the crops and other types of vegetation in the area. Other limitations of the data are that when some areas show up as clouds, then some of the surrounding region around the clouds may also have artificially low index values that should have been screened out. A third limitation is due to the compositing of 14 day intervals. Some of the pixel values could be as far as 14 days apart when ground conditions reflected could easily have been different conditions. NASS is looking forward to data from the new upcoming NASA sensor called MODIS. Because MODIS will have better spatial resolution, more bands, less atmospheric interference related problems, and still maintain the timely feature, we anticipate even better information on large area vegetative condition to be available.

V. CROP YIELD FORECASTING AND ESTIMATION

At present, there is a small joint research effort between NASS and USDA's Agricultural Research Service's Remote Sensing and Yield modeling laboratory. The effort is to examine the utility of remote sensing inputs to a yield model when combined with other inputs, such as plant process type models, weather data, soils data, AVHRR data, NASS's weekly crop stage and condition survey, and the crop specific categorized data from Landsat 5 TM described in the acreage section. This type modeling effort is being developed and tested currently for spring wheat in North Dakota. The current goal is to see if this procedure can provide another reliable input for final county level yield estimates in addition to the farmer reported survey data. To date, the results are not too encouraging, but we are continuing the effort as the new NASA sensor called MODIS should provide better measurements on plant condition, although still not crop specific. For different reasons, no type of spatial resolution of remotely sensed data is particularly promising for crop yield forecasting and estimation, when compared with conventional survey methods.

NASS's operational crop yield forecasting and estimation program makes no use of remotely sensed data or direct use of weather data. Survey data, farmer reported or objective yield sample survey data based on crop counts and measurements throughout the season are the major inputs to the Agricultural Statistics Board and the NASS State Statistical Offices in setting official yield forecasts and estimates at national, regional State and sub-State levels such as U.S. counties. A recent summary of the operational program methods is provided in the NASS publication "The Yield Forecasting Program of NASS" (July 1998).

VI. REFERENCES

Allen, J.D. and G.A. Hanuschak, 1988, "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.

Bellow, Michael and Mitchell Graham, 1992. "Improved Crop Area Estimation in the Mississippi Delta Region Using Landsat TM Data," American Society of Photogrammetry and Remote Sensing Convention, Washington, D.C.

Cotter, James and C. Mazur, 1992. "Automating the Development of Area Sampling Frame Using Digital Data Displayed on a Graphics Workstation," Eurostat Conference on New Technologies for Statistics, Bonn, Germany.

Cotter, Jim and Jack Nealon, 1987. "Area Frame Design for Agricultural Surveys," U.S. Department of Agriculture, National Agricultural Statistics Service.

Craig, Michael E., 1992. "Applications of Advanced Technology for Agricultural Statistics." United Nations, Conference of European Statisticians, Bratislave, Czech and Slovak Federal Republic.

FAO Statistical Development Series. Multiple Frame Agricultural Surveys, Vol. 1. 1996.

Ibid., Vol. 2, Ch. 4. 1998.

Hanuschak, G.A., R.D. Allen and W.H. Wigton, 1982. "Integration of Landsat Data into the Crop Estimation Program of USDA's Statistical Reporting Service 1972 - 1982." Paper presented at the 1982 Machine Processing of Remote Sensed Data Symposium, West Lafayette, IN.

Hanuschak, G.A., and K.M. Morrissey, 1977. "Pilot Study of the Potential Contributions of Landsat Data in the Construction of Area Sampling Frames," U.S. Department of Agriculture, Statistical Reporting Service.

Holko, Martin L., and Richard S. Sigman, 1984. "The Role of Landsat Data in Improving U.S. Crop Statistics." Paper presented at the Eighteenth International Symposium on Remote Sensing of Environment, Paris, France.

Houseman, Earl E., 1975. "Area Frame Sampling in Agriculture," U.S. Department of Agriculture, Washington, D.C.

Huddleston, Harold F., 1976. "A Training Course in Sampling Concepts for Agricultural Surveys," Statistical Reporting Service, U.S. Department of Agriculture, Washington, D.C.

Vogel, Frederick A. and Gerald A. Bange, 1999. "Understanding USDA Crop Forecasts," U.S. Department of Agriculture, National Agricultural Statistics Service and Office of the Chief Economist, World Agricultural Outlook Board.

"The Yield Forecasting Program of NASS," SMB Staff Report No. SMB-98-01, July 1998.