A Look at the Remote Sensing Applications Program of the National Agricultural Statistics Service

J. Donald Allen[1]

Feb. 1990

Abstract: This paper presents a summary of the procedures used by the National Agricultural Statistics Service (NASS) in its crop area estimation program during the years 1980–1987. It includes briefly some of the results with more detailed information provided by Allen and Hanuschak (1988). In its application program, NASS formed a regression estimator for crop acreage by using satellite data in conjunction with ground data which were collected during the annual June Enumerative Survey. The track record shows that the Landsat based crop area estimates for major producing regions of the U.S. were closer than the June Enumerative Survey (JES) direct expansion estimates to the Agricultural Statistics Board final estimates most of the time. The basic methodology, data processing techniques, and concepts used in the Landsat estimating program were developed through various research projects during the 1972–1979 period and are introduced briefly here. The timing of this description is appropriate

[1] Research and Applications Division, National Agricultural Statistics Service, Washington, D.C. 20250, U.S.A.

Acknowledgements: The author wishes to thank all of the NASS staff (headquarters and field offices) who worked in the area of remote sensing applications and research during the 1972–1987 period. A special thanks goes not only to those that are referenced in this paper but also to those who left files, notes, and letters behind as a trail for others to follow. This especially includes those managers involved in remote sensing applications and research who provided support over the years as well as costs summaries and documented project plans. We also wish to extend our appreciation to Sherri Finks, Suzy Babb and Amy Curkendall who diligently did the formatting and word processing of this paper. Our appreciation is also extended to the staffs of other participating organizations:

National Aeronatics and Space Administration
  Headquarters – Washington, D.C. (Pitt Thome – now with Destek Inc.)
  Ames Research Center (Buzz Slye)
  National Space Technology Laboratory
  Johnson Space Center (Gene Rice – now at NASA Headquarters – Space Station Project)

Boeing Corporation

Martin Marietta Data Systems (MMDS)

Bolt, Beranek and Newman Corporation (BBN)

U.S. Department of Agriculture's
  Soil Conservation Service
  Forest Service
  Agricultural Stabilization and Conservation Service (George Melvin)
  World Agricultural Outlook Board (Dick McArdle)
  Agricultural Research Service (Galen Hart)

California Department of Water Resources (Jay Baggett)

University of Houston (Dr. Raj Chhikara)
Purdue University (Dr. Marv Bauer – now with University of Minnesota)
University of California at Berkeley (Dr. Randy Thomas)
French Ministry of Agriculture (Jean Meyer-Roux – now with European Economic Community)
EOSAT Corporation (Dick Kott)
National Oceanic and Atmospheric Administration (NOAA)
U.S. Geological Survey (USGS)
EROS Data Center (Paul Severson)

with 1987 being the final year of operational crop estimation using data from the Landsat Multispectral Scanner Sensor (MSS). In the future there will most likely be a return to the use of satellite data in the estimating program, but for now NASS will no longer be using this data to produce timely crop estimates. The primary reason for this discontinuation is the uncertain status of the current Landsat satellites which have already outlived their expected design lives. In addition new satellite technology in the United States, France, Japan, India, and the USSR has produced data far superior to that yielded by Landsat's MSS. However, in order for NASS to take advantage of the new data, more research is required to assess the feasibility of its use so that when a new program is implemented, the anticipated improvement in the accuracy of the results will be cost effective.

Key words: Landsat; crop area estimates; satellite data.

1. Introduction

The National Agricultural Statistics Service (NASS), an agency within the United States Department of Agriculture (USDA), has the primary responsibility of providing statistics for domestic crop and livestock production. In particular, for major crops, estimates are made for planted and harvested acreage, yields, prices received, and, in some instances, stocks and disposition. For the most part, the statistics are derived from data collected through a variety of sample surveys. For principal crops, the major surveys include a national area frame survey in June; a quarterly multiple frame survey (list and area) in June, September, December, and March; and during the growing season, monthly objective yield surveys using actual field plots for yield forecasts. The area sampling frame used by the agency has been constructed and stratified based on land usage (primarily percent cultivated). NASS first began using remotely sensed data in the 1950s to aid in the construction of state area sampling frames; at that time, it was in the form of aerial photography. The use of earth resource satellite data from the U.S. Landsat was a natural extension in this process. In 1977, the value of photo-interpreting Landsat imagery in area frame construction was demonstrated (Hanuschak and Morrissey 1977).

Landsat's value as a digital input in the development of a viable crop estimator also became recognized. In concert with the launch of Landsat I in 1972, NASS statisticians were selected by the National Aeronautics and Space Administration (NASA) to be principal investigators on the use of Landsat data for agricultural statistics. With a small research staff, they proceeded to conduct a pilot test of combining conventional ground-gathered data with Landsat digital data to form a crop area estimator. The pilot was considered worth pursuing further with several years of research. Full state tests were conducted in Illinois in 1975 (Gleason, Hanuschak, Starbuck, and Sigman 1977) and Kansas in 1976 (Craig, Cardenas, and Sigman 1978). The first timely (i.e., end of season) crop area estimate was calculated for Iowa in 1978 (Hanuschak et al. 1979). This experience and NASS's participation and evaluation role in the governmental LACIE (Large Area Crop Inventory Experiment) project in the mid-1970s led NASS to the point of large scale applications.

The progression of Landsat data usage in the crop estimating program of NASS was given additional impetus by the initiation of the AgRISTARS (Agriculture and Resource Inventory through Aerospace Remote Sensing) program which began on October 1, 1979. Initially, this was to be a six-year project set to end September 30, 1985, but was later extended to October 1,

1986. It was an interagency program involving not only the USDA, but also the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration, the U.S. Department of Interior, and the Agency for International Development. Its focus was on the possible uses of aerospace remotely sensed data to answer agricultural resource questions as well as to meet identified information needs of the USDA. Moreover, it was to determine the usefulness, cost, and extent to which these data could be integrated into existing or future USDA systems so as to improve the objectivity, reliability, timeliness, and adequacy of information required to carry out USDA missions (NASA 1981).

One of the major initiatives under AgRISTARS was to apply the above objectives toward estimating domestic crops and land covers. It was in this area that the NASS research staff took the leadership role (NASA 1981). Plans were to make crop estimates in two states in 1980 using the Landsat data, and each succeeding year, two more states would be added with the goal of having ten states in the estimating program by 1984. The crops to be estimated were to be large area crops in homogeneous regions. In addition, the intent was for research to continue in the area of determining land covers. The ultimate goal of the Landsat estimation program was to provide timely estimates of crop acreages which would have significantly smaller sampling errors than the estimates that were then in use.

In summary, NASS employed a four-prong approach in using the satellite data: (1) remote sensing was to be viewed as just another method of data collection, (2) remote sensing would be used to supplement existing efforts, (3) integration into the estimation program would be founded on strong statistical procedures, and (4) resource effective techniques would have to be developed for the program to be successful. It was realized from the start that there would be both challenges and some constraints in using satellite data. First, it was understood that it would be used as auxiliary data since statistical defensibility would require some source of ground data to be used to insure proper categorization of the satellite information. Also, it was realized that cloud cover problems could prevent some of the data from being available during the times when it might be needed. There would also be massive amounts of data to be processed which might limit the rate of the program's expansion. Additionally, completion of the crop acreage estimates would come at the end of the crop year because of the time needed to perform the analysis; this meant that the figures could not be used in early season forecasts. Finally, the resolution (i.e., degree of interpretability) of the MSS data would be such that some crops and some states with small field sizes could not be included in the estimation program.

2. The Landsat Space Program

NASA's Landsat satellite series began with the launch of Landsat I in July 1972. This was followed by Landsat II in January 1975 and Landsat III in March 1978. The first two spacecrafts were equipped with Return Beam Vidicon (RBV) cameras and Multispectral Scanners (MSS) while Landsat III provided data from an High Resolution Panchromatic (also referred to as RBV) camera as well as from an MSS. The decision was made to use the MSS data as opposed to the RBV data since MSS was in a form that was more adaptable for computer processing. It supplied four bands of data for analysis and a spatial resolution of eighty square meters (later sixty meters). Resolution in this context can be thought of as the ability of the imaging system to distinguish closely spaced objects in the subject area; in essence, this means that the spatial resolution can be thought of as the minimal ground area in which the sensor is sensitive to radiation. A more detailed discussion of the satellite data used can be found in Appendix A. The point to point fly over period for these crafts was eighteen days. Landsat IV was launched in July 1982 and Landsat V in March 1984. The point to point fly over period was sixteen days for these two satellites. They were equipped with the Multispectral Scanners as well as Thematic Mappers (TM). The TM data was considered superior since it provided seven bands of data as well as thirty-meter resolution. However, all previous research by NASS had been directed as MSS data; in addition, the TM data was roughly five times more expensive to buy than MSS and even more expensive to process. These factors influenced the agency's decision to continue with MSS data until the benefits of the Thematic Mapper could be investigated further.

Due to vastly improving computer technology, the agency is currently in the process of researching the use of TM data. In addition, data from the French SPOT satellite which was launched in 1986 is also being examined. The French spacecraft provides three bands of MSS type data but with twenty-meter resolution and a ten-meter panchromatic band. The French SPOT satellite is also pointable which is a definite advantage for maximizing the probability of cloud free coverage. There are possibilities for other data since Japan and India also have recently launched satellites with capabilities similar to those of the Landsat series. In addition, there are plans at the present to introduce Landsat VI in 1991. The intentions are that this Landsat will be equipped with a Thematic Mapper but not a Multispectral Scanner. As a result, data in the format currently used by NASS will no longer be available from sources in the United States once Landsat IV and V fail. At this time, both are deteriorating with only one satellite providing MSS data and the other providing only TM data.

There are normally only three different Landsat products used by NASS's statisticians in their analysis process. The first of these are 1 : 1,000,000 scale transparencies. These are used to evaluate cloud coverage and, in turn, to decide what combinations of imagery dates can be used to provide the best data. Secondly, there are 1 : 250,000 scale black and white paper products which are actually photographic interpretations of the scenes. These are used in the registration process (see Section 3). Lastly, there are the data tapes themselves. The costs of these products remained relatively stable up to 1983 when a large price increase occurred. Expenditures in this area have ranged annually from 10 to 15 percent of the total project costs since 1983.

3. Methodology

The methodology used by NASS in its crop estimating programs is best described in U.S. Department of Agriculture (1983). Currently, major surveys use multiple sampling frames. More precisely, a list frame is used from which samples are drawn with an area frame used to account for the incompleteness which is inherent in the list. It should be noted that the area frame can stand alone as a complete frame covering the entire population and, as a result, is used to provide indications for crop acreages as well as being a complementary part of a multiple frame indication. The area frame estimator works well for major crops and livestock inventories, but not so well with minor or rare items such as specialized crops. Area frame procedures have been most recently described by Cotter and Nealon (1987).

A major source of crop data for NASS is the June Enumerative Survey (JES) which is conducted nationwide each year. This particular survey relies only on the area frame which is itself a stratified population. From the frame, a sample of approximately 16,000 land segments is randomly selected with the segment size ranging from 40 to 2400 acres and averaging 450 acres. The ground data are expanded to state, regional, and national totals using methods similar to those outlined in Cochran (1977) for stratified sampling. At the national level, sampling errors are normally less than two percent for major crops such as corn, soybeans, and winter wheat. Survey training programs, efforts at standardization, and edits and analysis are used to reduce and control the nonsampling errors. Crop acreage estimates based on the JES are published around July 10 each year. Final year end estimates are published around January 15 of the following year.

In crop acreage estimation, the Landsat data are used in conjunction with the JES data in the form of a regression estimator. The exact nature of this estimator as used by NASS was initially described by Von Steen and Wigton (1976). The estimator also has been recounted in a number of other NASS research reports, and most recently it was described in detail by Holko and Sigman (1984). Studies have cited two technical problems with the regression estimator: (1) a bias may exist in the regression estimator when sample sizes are small (Chhikara and McKeon 1985; Lundgren 1984) and (2) a bias may exist as a result of using the same area frame segments to estimate both the parameters of the discriminant functions and the regression equation (Jones 1987; Holko 1984; Zuttermeister 1985; Gleason et al. 1977). Current and future research projects will address these issues.

Normally, it takes approximately three to four months to obtain and analyze fully the Landsat data for a state. Also, it is desirable that the data being used pertain to the optimum growing period for the crop being estimated. For winter wheat in the central part of the United States, this means that Landsat data would normally be at their best if they related to the period between mid-April and the end of May. The optimum for the spring crops being estimated would ideally relate to the period between mid-July and the end of August. Because of this timing, the Landsat indications are used only in setting end of the season acreage estimates.

Briefly, the process begins with the calibration of the JES land segments to a map base; that is, the exact location of a segment is translated into a set of latitudinal and longitudinal coordinates. Then, the ground data are collected through the JES, edited, and put into machine language. The field boundaries that were indicated during the JES are then digitized along with the segment boundaries. Each frame or scene of Landsat data, each covering an area of 170 kilometers by 185 kilometers, must also be registered or assigned latitudes and longitudes. The next step requires that the segments and fields be mapped on to the Landsat scenes using the coordinate system that was derived during calibration and registration. Each pixel (a square area covering 0.8 acres of a scene) that overlaps a JES segment is assigned a crop or cover type based on the corresponding JES ground data. In addition, each pixel also has a set of MSS measurements. In the ensuing phase, a clustering algorithm is applied to the set of pixels representing each crop. Each of the resulting clusters has associated with it a mean vector and a covariance matrix (i.e., a signature).

All the segment data are then classified into crop types based on clustering results. Next a regression relationship is developed between the ground data and the classified pixels. The entire Landsat scene is then classified based on the clustering of the sample ground data. In the succeeding step, the regression relationship is applied to the full scenes. All the data are finally aggregated across scenes from the Landsat estimate. The final estimate will also include JES expansions for areas not covered by Landsat scenes. Domain estimators are used in these situations (Cochran 1977).

In more specific terms, the JES yields a direct expansion estimate for crop acreage which is based solely on the JES data while the regression estimator uses the JES data and the Landsat data.

3.1. *JES direct expansion*

For a given state, let $h = 1, 2, \ldots, L$ denote the land use strata. Within each stratum, the total land area contains N_h primary sampling units from which n_h units (segments) are selected. Using only the area frame data collected during the JES, the direct expansion estimator for the total acreage of a particular crop in any given state can be expressed as

$$\hat{Y} = \sum_{h=1}^{L} N_h \bar{y}_h$$

where $\bar{y}_h = \Sigma_{j=1}^{n_h} y_{hj}/n_h$ and y_{hj} is the reported acreage of the specified crop in segment j in stratum h. The corresponding variance estimate is given by

$$\mathrm{Var}(\hat{Y}) = \sum_{h=1}^{L} \frac{N_h^2}{n_h(n_h - 1)} \frac{N_h - n_h}{N_h} \times \sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2.$$

3.2. *Regression estimation*

The total based on the regression estimator is given by

$$\hat{Y} = \sum_{h=1}^{L} N_h \bar{y}_{h(\mathrm{reg})}$$

where $\bar{y}_{h(\mathrm{reg})} = \bar{y}_h + \hat{b}_h(\bar{X}_h - \bar{x}_h)$ and $\hat{b}_h$ is the estimated regression coefficient for land use stratum h based on regressing ground reported acres on classified pixels in the n_h sampled segments.

Here $\bar{X}_h$ is the average number of pixels classified to the specified crop in each PSU; that is, all frame units are included in the calculation and $\bar{X}_h = \Sigma_{i=1}^{N_h} X_{hi}/N_h$ where X_{hi} is the number of pixels in the ith frame unit of stratum h. Similarly, $\bar{x}_h$ is the average number of pixels classified to the crop in each of the sampled segments in stratum h; that is, only the sampled frame units are included and $\bar{x}_h = \Sigma_{j=1}^{n_h} x_{hj}/n_h$ where x_{hj} is the number of pixels in the jth sample unit of stratum h. The corresponding variance estimate is given by

$$\mathrm{Var}(\hat{Y}) = \sum_{h=1}^{L} \frac{N_h^2}{n_h} \frac{N_h - n_h}{N_h} \times \sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2 \frac{(1 - \hat{R}_h^2)}{n_h - 2} \times \left(1 + \frac{1}{n_h - 3}\right)$$

where $\hat{R}_h^2$ is the coefficient of determination between the reported acreage for the specified crop and the corresponding pixels classified to the crop. Note that the variance of the regression estimator approaches zero as $\hat{R}_h^2$ approaches unity for fixed n_h. In other words, as the correlation increases between the ground data and the classified Landsat data, the variance in the regression estimator decreases.

Since Landsat data cannot be obtained for an entire state on any given date, states are partitioned into separate analysis areas or districts for each date for which Landsat data are obtained. The regression estimate is developed within each analysis district. As referred to earlier, the statistical characteristics for the set of pixels in the sample must first be developed. Initially, all pixels of a known crop type are grouped together and then clusters are formed. Pixels forming field or segment boundaries as well as those for which ground data are felt to be inadequate are excluded. Again each analysis district is processed separately. There are two clustering algorithms which can be used. The first algorithm relies on the ISODATA algorithm and is referred to as "ordinary clustering" (Ball and Hall 1967); it works suitably for small data sets but the processing costs incurred with its use make it less suitable for large data sets. The second is Classy which is a maximum likelihood clustering algorithm developed at Johnson Space Center in Houston, Texas (Lennington and Rassbach 1979); it functions best if there are a large number of pixels (greater than 500). Additionally, Classy relies more on normality assumptions than does ordinary clustering. One or more clusters result for each land cover. Statistics surrounding the resulting clusters are examined which in effect allows the analyst to perform editing functions to improve the accuracy of the classification. Among the statistics provided are two measures of separability: Swain-Fu (Swain 1972) and transformed divergence (Swain and Davis 1978).

Once the signatures for the clusters have been determined, all the pixels in the sampled segments within an analysis district are classified using this information. This is performed separately for each analysis district. Multivariate normality is assumed which seems to be justifiable for most remote sensing applications. Additionally, classification accuracy seems to be robust to violations of the assumption (Swain and Davis 1978). In particular, quadratic discriminant scores based on multivariate normality are calculated for each pixel in each analysis district

$$d_i^Q(\mathbf{x}) = -0.5 \ln \{\det \mathbf{S}_i\} - 0.5(\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{S}_i^{-1}(\mathbf{x} - \bar{\mathbf{x}}_i) + \ln p_i$$

where $i = 1, 2, \ldots, g$ represents the individual clusters. In the formulation, p_i represents the prior probability that a pixel belongs to population i and $\mathbf{S}_i$ is the sample covariance matrix. So for each pixel, g discriminant scores are computed. A pixel then would be assigned to population k if

$$d_k^Q(\mathbf{x}) = \text{largest of } d_1^Q(\mathbf{x}), \ldots, d_g^Q(\mathbf{x}).$$

The classification is performed with equal prior probabilities as well as with distinct priors. There are several ways to derive the values for unequal probabilities. The most common procedure is to assume that the ground data are completely "true" and then to compute the number of pixels corresponding to each ground cover; a weight reflecting the resulting proportions is subsequently assigned to each land cover category. Since each land cover category may consist of several clusters, this weight is further proportioned to reflect the percentage of pixels that are in each cluster within a category. Several sets of classifiers are used up to the point of the final accumulation of analysis district estimates to a state level estimate. That way, the analyst has the opportunity to continue the assessment of each set's performance up to the final step of the process.

Each set of classifiers is applied separately to all the pixels in each analysis district. The results of this full frame classification as well as the results from the sampled segment classification are then used as inputs into the regression estimator that was outlined earlier. The final step of the process calls for combining all the estimates which at this point are at an analysis district level.

Typically, a classifier is evaluated by examining its confusion matrix. However, if a classifier is to be used to estimate crop acreage, its performance should be assessed on that basis. This means that the variance of the regression estimate can be thought of as the measure of the classification's adequacy (Gleason et al. 1977). The best estimate in this sense is obtained by maximizing the correlation between the ground data and satellite data. The relationship between the two can be affected by numerous factors. Among these are the region of interest, the data of the imagery (since there exists an optimum time frame for a given crop in a given region), and the number of pixels used in developing the clusters. The correlations are also affected by the prior probabilities used, the number of clusters decided upon as well as the number of land cover categories, and whether or not some land use strata are pooled or excluded during the analysis. The decision to exclude segments is usually made in instances where there are less than five segments within a given strata within a particular analysis district while pooling usually is done only for strata which have the same expansion factors. The average field size was also found to have an effect with results better for those crops grown in larger fields (Cardenas and Hanuschak 1978); for this reason, the methodology is best suited for major crops in the principal producing areas. Additionally, the satellite data themselves are assigned a quality level by the data providers; obviously, the better the quality of the data, the higher one's expectations would be for the accuracy of the classification process.

One of the major problems which must be dealt with in using Landsat data is cloud cover. Each year every effort was made to obtain cloud free imagery within the optimum period, but experience showed this to be an impossible task. Therefore, it became necessary to address the issue in the acreage estimates. A post stratification approach is used for this purpose. Areas for which Landsat data are available compose one post stratum; here the regression estimator as outlined earlier is used. For areas not covered by Landsat data, the direct expansion estimator is applied. This approach is possible since the total number of frame units as well as the sampled units within areas covered and not covered are known.

In order to determine the success associated with the regression estimator, its relative efficiency (RE) is calculated. The RE is a measurement of the gain in precision of the regression estimator as compared to the JES direct expansion

$$\text{RE} = \frac{\text{Variance (JES direct area expansion)}}{\text{Variance (Landsat/JES combined regression estimator)}}.$$

Here the variance for the JES direct expansion is based on the June survey alone and does not reflect any updates in the data that resulted from follow-up contacts necessitated by the remote sensing project. In general, the relative efficiency can be thought of as the factor by which the JES sample size would have to be increased in order to yield a direct expansion with a variance equal to that obtained using the Landsat data. The approximate breakeven RE in terms of cost effectiveness was calculated to be approximately in the 2.5 to 3.5 range based on 1981 data (Hanuschak, Allen, and Wigton 1982). That is, an RE above that range would indicate that the same precision could be obtained at less cost if the Landsat-JES approach was used as opposed to expanding the JES sample size.

The cost-benefit ratio is certainly not an exact measure and is subject to many assumptions and questions, such as:

1. Would or would not the agency continue a general purpose JES?
2. What would be the financial benefit to the agricultural sector of the economy if the corn acreage variance estimate was cut by a factor of two to three?
3. What would be the cost of a ground survey for only crop and land use information for remote sensing analysis and not a general purpose JES?
4. What is the value of county level remote sensing estimates that the JES does not provide?
5. What would be the magnitude of the gains in estimating other items if the sample size of the JES were to be doubled or tripled?
6. Could the general purpose JES be doubled or tripled in terms of overall response burden and implementation (enumerator and state office workload, potential nonsampling errors, etc.)?

Over time, the costs of the JES have increased while the remote sensing costs have decreased. As a result, the breakeven point for cost effectiveness has declined considerably. In 1987, by the same criterion, relative efficiencies exceeding the 1.5 to 2.5 range would be considered an indication of cost effectiveness. A graphical representation of the total Landsat costs and the JES costs follow in Figure 1. The Landsat costs are all costs associated with the Landsat estimate above and beyond the operational JES costs.

4. Program Coverage and Results

From 1980 to 1987, the program expanded from two states (Iowa and Kansas) and three crops (corn, soybeans, and wheat) to eight states (Arkansas, Colorado, Illinois, Indiana, Iowa, Kansas, Missouri, and Oklahoma) and three additional crops (cotton, rice, and sorghum). This alone was a significant accomplishment with nearly the same personnel resources. The growth of

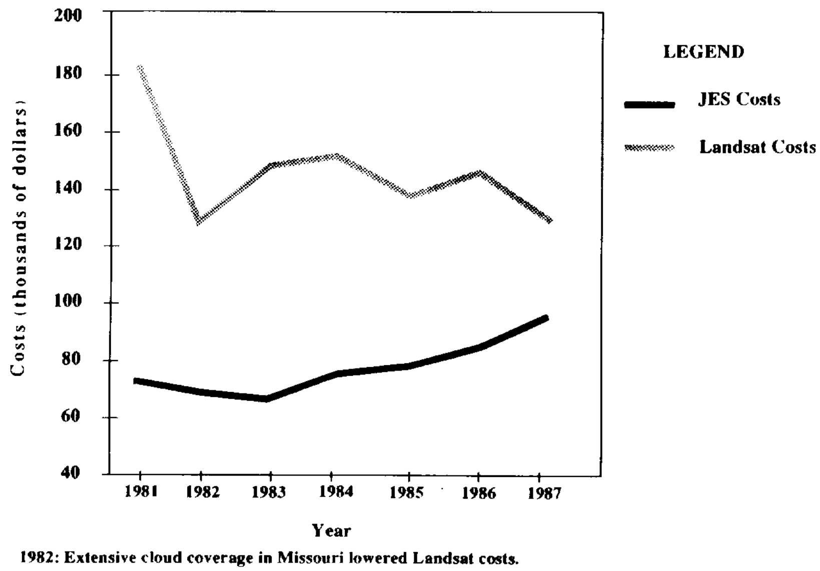

Fig. 1. Average costs per state in Landsat program Landsat versus JES

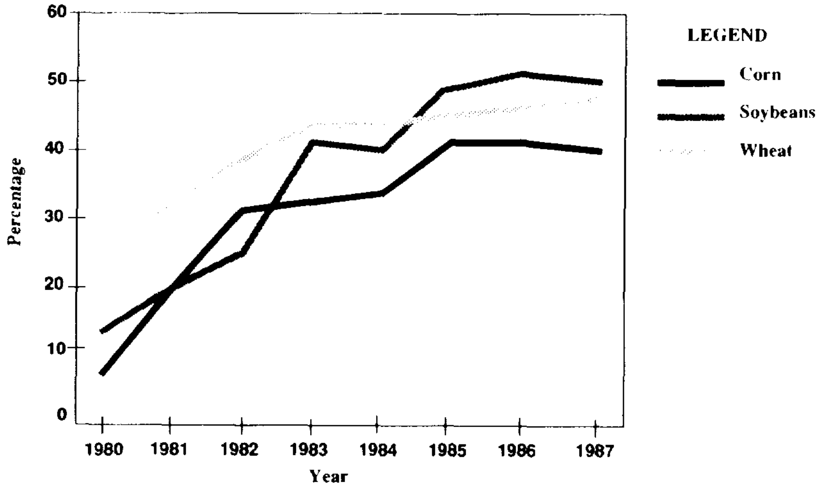

Fig. 2. Percent of United States planted acreage accounted for by states in Landsat program

the project is reflected in Figure 2. Across all crops and all years there were thirty-nine regional estimates calculated. Out of these thirty-nine estimates, the Landsat based estimator was closer to the Agricultural Statistics Board's final estimate twenty-one times (see Table 1). While this is not a statistically significant difference, it is still of interest to see such a new and complex technology outperform the June Enumerative Survey, in terms of estimated accuracy. The comparison assumes the Board's final estimate to be the best approximation to "true values." Comparisons of state estimates can be found in Allen and Hanuschak (1988).

5. Project Cost

Project cost data were diligently maintained during the eight-year stint of the remote sensing estimation program. One's perspective of these costs can be enhanced somewhat by reviewing the historical cost data for various remote sensing research efforts conducted prior to 1980. When the first full state project (Illinois – 1975 data) was conducted, the total cost was $750,000. The project encompassed two years and included much of the developmental phase of EDITOR (the computer software used for the analysis). The first real time full state application project (Iowa – 1978 data) by comparison cost $300,000. The first actual operational year (1980) costs were $200,000 per state; this was reduced to an average of approximately $140,000 per state for the 1982–1986 period and ended with an average of $129,000 per state for 1987. None of the cost data have been adjusted for inflation which would show an even more dramatic cost reduction over time. The two major reasons for the sharp drop in costs were (1) the efficient use of computer resources with a range of applications involving supercomputers, mainframe, mini-, and microcomputers, and (2) the inclusion of more states in the project while maintaining approximately the same level of personnel. Additionally, the proper mix of hardware and optimized software tended to maximize the gains in cost efficiency. The graph that follows (Figure 3) shows detailed costs for 1981–1987.

Table 1. Comparison of final Agricultural Statistics Board's (ASB) final estimate with June Enumerative Survey (JES) and Landsat based estimates

Year	CORN PLANTED ACRES		SOYBEAN PLANTED ACRES	
	JES Estimate as percent of ASB Final	Landsat Estimate as percent of ASB Final	JES Estimate as percent of ASB Final	Landsat Estimate as percent of ASB Final
1980	98.4	99.5*	100.8*	98.0
1981	100.3*	99.8	104.0	97.9*
1982	101.1	99.5*	103.3	100.4*
1983	103.4	99.8*	100.7*	97.9
1984	99.8*	97.1	103.4	99.3*
1985	100.1*	99.4	100.9*	98.5
1986	98.6	99.1*	103.7	101.3*
1987	100.0*	97.6	102.6	101.2*
Average	100.2*	98.9	102.4	99.3*

Year	WHEAT HARVESTED ACRES		RICE PLANTED ACRES	
	JES Estimate as percent of ASB Final	Landsat Estimate as percent of ASB Final	JES Estimate as percent of ASB Final	Landsat Estimate as percent of ASB Final
1980	107.4	104.0*	–	–
1981	107.7	103.9*	150.6	100.0*
1982	106.4	101.4*	–	–
1983	104.0	100.9*	117.4	109.2*
1984	101.0*	98.1	97.0*	96.7
1985	103.0	100.0*	102.7*	109.7
1986	102.1*	95.7	91.2	94.3*
1987	100.6*	96.0	90.3*	88.1
Average	104.1	100.0*	108.2	99.7*

Year	COTTON PLANTED ACRES		SORGHUM PLANTED ACRES	
	JES Estimate as percent of ASB Final	Landsat Estimate as percent of ASB Final	JES Estimate as percent of ASB Final	Landsat Estimate as percent of ASB Final
1983	98.4*	76.9	–	–
1984	89.0	104.4*	105.8	95.0*
1985	93.0*	109.4	102.4*	89.9
1986	133.9	120.1*	96.3*	86.3
1987	122.1	108.1*	99.4*	93.1
Average	107.3	103.8*	101.0*	91.2

*Most accurate result compared to the ASB final.

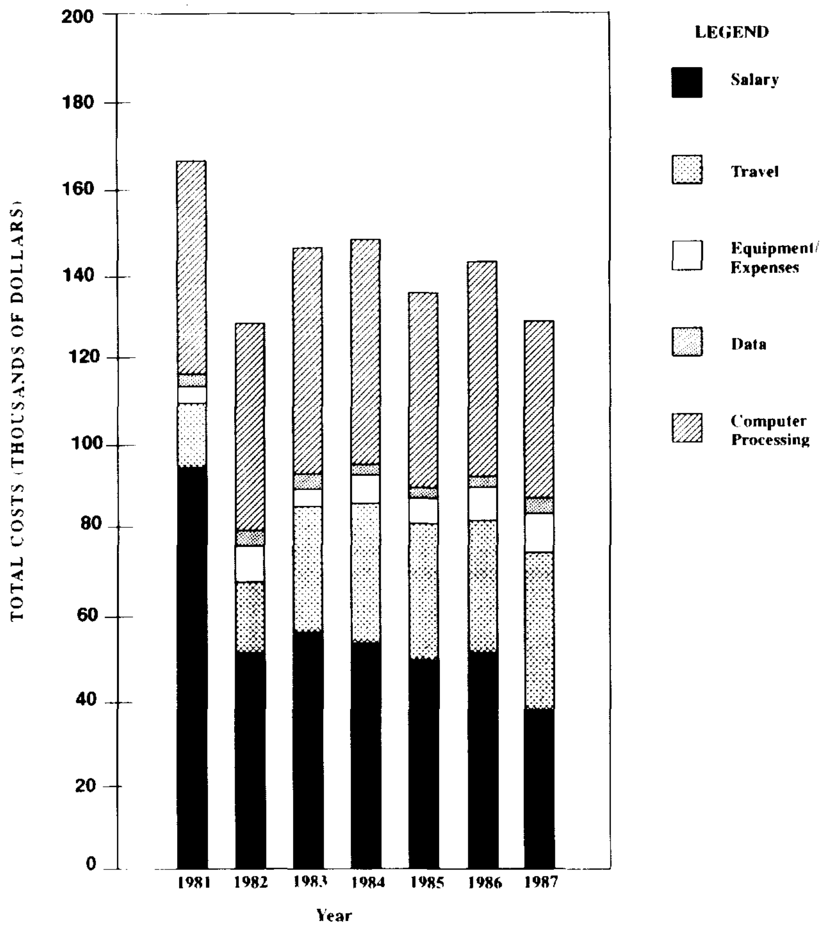

Fig. 3. Average costs per state for Landsat estimation program: 1981–1987

6. Recent Developments

At the end of 1987, the decision was made to rechannel some of the resources of the remote sensing applications group back into a research program. There were several reasons for this. Foremost was the concern over the anticipated failure of the current Landsat satellites. At the present time, Landsat IV and V are still functioning. However, the life expectancies of these satellites are elapsing. The failure of either would create obstacles for NASS. Past studies have shown that two MSS satellites are needed in order to provide coverage adequate to overcome cloud coverage (Winings 1982), and currently, all of the methodology is geared toward the use of the MSS data so the use of other types of data is not possible at this time. This is further compounded by the fact that future Landsat satellites will not carry multispectral scanners. In addition, there is a need to study the new high resolution sensors on the pointable French Spot satellite as well as Landsat's Thematic Mapper to evaluate their suitability for use in an operational program. This all translates into the need for more very well targeted research. Another factor influencing the decision to defer the applications program was the reduction in available funding. Federal budget cuts have demanded that some projects be curtailed, and remote sensing applications with expenditures exceeding one million dollars per year was management's choice when coupled with the above considerations. On a positive note, the cessation of the remote sensing applications program is being viewed as only a transition period that will lead to an even better course of action in the use of advanced higher resolution satellite data. The advanced sensor data from the commercial systems of the 1990s will contain improved spatial, spectral, and temporal information in the data. To NASS, this should translate into more accurate acreage estimates and perhaps also crop condition or yield assessments if expenses can be controlled through cost effective methodology. Additionally, the program change also has allowed for more resources to be applied to research on computer assisted area frame construction. In connection with this project, NASS has recently received a grant from NASA headquarters in support of this endeavor.

7. Conclusions and Observations

During the eight-year project, there were three major findings. First, crop area estimates using Landsat data for large areas could be produced in a timely fashion with relatively small additional staffing. Secondly, the Landsat based estimates had considerably lower sampling errors than the JES ground survey, and the regional Landsat estimates tended to be closer to the Agricultural Statistics Board final estimates. Third, based on internal agency costs and benefits, the extra cost of processing the Landsat data were near a breakeven point for corn, soybeans, wheat, and sorghum and above breakeven for cotton and rice. Cloud cover, technical problems with the satellites and ground systems, and the limits on the amount of information contained in Landsat's MSS data seemed to be the major problems encountered that prevented more positive comparisons. Additionally, there are many other considerations, findings, and benefits from this eight-year project. It is not feasible to recall or list each and every one of them but some of those which potentially could be termed as having major significance include the following:

1. The agency research staff gained considerable experience in the use of supercomputers.
2. The agency research staff gained considerable experience in the use of specialized hardware for digitization, both vector and video. The knowledge gained in vector digitization and in the visual interpretation of Landsat imagery in terms of land use was conveyed to the area sampling frame construction staff and has subsequently paid large dividends in the efficiency of area frame construction as well as in related quality control.
3. The agency now has a small highly trained staff to evaluate the more advanced satellite sensors of today and the future.
4. The agency staff gained an international reputation for its Landsat methodology and large scale inventory capabilities as well as its efficient use of supercomputers.
5. Statistical formulas were developed for small area (county level) crop acreage estimates (Amis, Lennington, Martin, McGuire, and Shen 1982; Battese and Fuller 1981; Cardenas, Blanchard, and Craig 1978; Huddleston and Ray 1976; Walker and Sigman 1982; Chhikara and McKeon 1987).
6. The research staff of NASS, along with

the USDA's Soil Conservation Service and several other federal and state government agencies, demonstrated that land cover inventories in addition to crop acreage inventories could be successfully estimated using the Landsat regression procedures.

7. The use of Landsat data for yield forecasting and estimation was examined but it was determined that any increase in information on crop yields was not cost effective compared to NASS's conventional yield surveys.
8. To the author's knowledge, economic benefit studies to determine the "value of the agricultural economy" of more accurate crop area estimates have not been conducted for the 1980–1987 period by professional economists. Therefore, thorough results from a cost-benefit analysis (considering both internal and external factors) are not available.

Appendix A: Discussion of Satellite Data

The Landsat data used in NASS's remote sensing program during the 1980–1987 period consisted of a set of measurements made by the satellite's multispectral scanner (MSS). Each measurement for the first three satellites in the series encompassed an area of 1.0 acres while the area was 0.8 acres for the other two. This measurement area is referred to as a pixel. Each of the satellites in the Landsat series was designed to travel in a nearly circular polar orbit with 14 orbits per day. Repeated coverage occurred every 18 days for Landsat I, II, and III while the repetition was every 16 days for the later versions. The orbiting paths covered widths of 185 kilometers and provided cross sections that were 170 kilometers in length. The scanner itself is a camera like device that divides the image being received into pixels and then measures the brightness of each pixel along the electromagnetic spectrum. Specifically, the total radiance of an object is measured in four bands of the spectrum. Two are in the visible portion of the spectrum (0.5–0.6 micrometers and 0.6–0.7 micrometers) while the other two are in the near infrared portion (0.7–0.8 micrometers and 0.8–1.1 micrometers). These four measurements provide a spectral "signature" for an object. The differences in these signatures allow for the classification of the pixels. The spatial resolution in the early satellite versions was 80 meters square but this was reduced to 60 meters for Landsat IV and V.

Thematic Mappers (TM) were provided on Landsats IV and V and were advanced sensors compared to the Multispectral Scanners. These devices operate in seven spectral bands with 30 meter resolution. Three of the bands are visible with an overall range of 0.45 micrometers to 0.69 micrometers. There is one near infrared (0.76–0.90 micrometers) and two shortwave infrared bands (1.55–1.75 micrometers and 2.08–2.35 micrometers). The final band is thermal infrared (10.50–12.50 micrometers). Measures for all but the latter are provided with 30 meter spatial resolution with the thermal band having 120 meter resolution. Pixel size for TM data is 0.2 acres (U.S. Geological Survey and National Oceanic and Atmospheric Administration 1984).

The first SPOT (Satellite Pour l'Observation de la Terre) satellite which was launched in 1986 follows a near polar orbit similar to the Landsat satellites. There are two sensors on board known as HRVs (high resolution visible). They produce multispectral images with 20 meter spatial resolution with measurements on two visible bands (0.50–0.59 micrometers and 0.61–0.68 micrometers) as well as a single near infrared band (0.79–0.89 micrometers). Additionally, black and white images, using the spectral band ranging from 0.51 to 0.73

micrometers, are produced with 10 meter resolution. The resulting pixel sizes are 0.098 acres and 0.025 acres respectively. Orbits repeat every 26 days. However, since the sensors are pointable, the same ground area can be observed several days in a row. The path width is 60 kilometers which can be adjusted to 81 kilometers. The cross sectional lengths are a constant 60 kilometers. Plans call for launching SPOT-2 some time in 1989 and SPOT-3 in the early 1990s (SPOT Image Corporation 1988).

8. References

Allen, J.D. and Hanuschak, G.A. (1988). The Remote Sensing Applications Program of the National Agricultural Statistics Service 1980–1987. National Agricultural Statistics Service staff report no. SRB-88-08, United States Department of Agriculture, National Agricultural Statistics Service.

Amis, M.L., Lennington, R.K., Martin, M.V., McGuire, W.G., and Shen, S. S. (1982). Evaluation of Small Area Crop Estimation Techniques Using Landsat and Ground-Derived Data. Report no. LEMSCO-17597, Lockheed Engineering and Management Services Company, Inc.

Ball, G.H. and Hall, D.J. (1967). A Clustering Technique for Summarizing Multivariate Data. Behavioral Science, 12, 153–155.

Battese, G.E. and Fuller, W.A. (1981). Prediction of County Crop Areas Using Satellite Data. Proceedings of the Section on Survey Research Methods, American Statistical Association, 500–505.

Cardenas, M., Blanchard, M., and Craig, M. (1978). On the Development of Small Area Estimators Using Landsat Data as Auxiliary Information. Staff report, United States Department of Agriculture; Economics, Statistics, and Cooperative Service.

Cardenas, M. and Hanuschak, G.A. (1978). An Exploratory Study on the Effects of Field Size and Boundary Pixels on Crop Signatures. Internal working paper, United States Department of Agriculture; Economics, Statistics, and Cooperative Service.

Chhikara, R.S. and McKeon, J.J. (1987). Estimation of County Crop Acreages Using Landsat Data as Auxiliary Information. Houston, Texas: University of Houston.

Chhikara, R.S. and McKeon, J.J. (1990). Estimation of Finite Population Under a Calibration Model. Journal of Official Statistics, 6, 295–310.

Cochran, W.G. (1977). Sampling Techniques. New York: John Wiley.

Cotter, J. and Nealon, J. (1987). Area Frame Design for Agricultural Surveys. Staff report, United States Department of Agriculture, National Agricultural Statistics Service.

Craig, M.E., Cardenas, M., and Sigman, S.R. (1978). Area Estimates by Landsat Kansas 1976 Winter Wheat. Staff report, United States Department of Agriculture, Economics Statistical Service.

Gleason, C.P., Hanuschak, G.A., Starbuck, R.R., and Sigman, R.S. (1977). Stratified Acreage Estimates in the Illinois Crop-Acreage Experiment. Staff report, United States Department of Agriculture, Statistical Research Service.

Hanuschak, G.A. and Morrissey, K.M. (1977). Pilot Study of the Potential Contributions of Landsat Data in the Construction of Area Sampling Frames. Staff report, United States Department of Agriculture, Statistical Research Service.

Hanuschak, G.A., Craig, M., Ozga, M., Luebbe, R., Cook, P., Kleweno, D., and Miller, C. (1979). Obtaining Timely Crop Area Estimates Using Ground-Gathered and Landsat Data. Technical bulletin no.

1609, United States Department of Agriculture; Economics, Statistical, and Cooperative Service.

Hanuschak, G.A., Allen, R.D., and Wigton, W.H. (1982). Integration of Landsat Data Into the Crop Estimation Program of United States Department of Agriculture's Statistical Reporting Service 1972–1982. Paper presented at the 1982 Machine Processing of Remote Sensed Data Symposium, West Lafayette, Indiana.

Holko, M.L. and Sigman, R.S. (1984). The Role of Landsat Data in Improving United States Crop Statistics. Paper presented at the Eighteenth International Symposium on Remote Sensing of Environment, Paris, France, October 1–5, 1984.

Holko, M. (1984). A Remote Sensing Application Methodology and Results. Unpublished working paper, United States Department of Agriculture, Statistical Research Service.

Huddleston, H.F. and Ray, R. (1976). A New Approach to Small Area Crop Acreage Estimation. Paper presented at the Annual Meeting of the American Agricultural Economics Association, State College, Pennsylvania.

Jones, E.M. Jr. (1987). Classifier Study. Unpublished manuscript, United States Department of Agriculture, National Agricultural Statistics Service.

Lennington, R.K. and Rassbach, M.E. (1979). Mathematical Description and Program Documentation for CLASSY – An Adaptive Maximum Likelihood Clustering Method. Report no. LEC-12177 (JSC-14621), Lockheed Electronics Company, Inc.

Lundgren, J.C. (1984). A Study of Bias and Variance in Landsat Data Based Regression Estimates for Crop Surveys Using Simulated Data. Report no. LEMSCO-20239, Lockheed Engineering and Management Services Company, Inc.

NASA (1981). AgRISTARS, Annual Report – Fiscal Year 1980. AP-JO-04111, Houston, Texas: NASA.

SPOT Image Corporation (1988). SPOT User's Handbook (vol. 1). Reston, Virginia: Spot Image.

Swain, P.H. (1972). Pattern Recognition. A Basis for Remote Sensing Analysis. LARS information note 11572, West Lafayette, Indiana, Purdue University.

Swain, P.H. and Davis, S.M. (1978). Remote Sensing. The Quantitative Approach. New York: McGraw-Hill.

United States Department of Agriculture, Statistical Reporting Service (1983). Scope and Methods of the Statistical Reporting Service. Miscellaneous publication no. 1308, Washington, D.C.: United States Department of Agriculture.

United States Geological Survey and National Oceanic and Atmospheric Administration (1984). Landsat 4 Data Users Handbook. Alexandria, Virginia: United States Geological Survey and National Oceanic and Atmospheric Administration.

Von Steen, D. and Wigton, W. (1976). Crop Identification and Acreage Measurement Utilizing Landsat Imagery. United States Department of Agriculture, Statistical Research Service.

Walker, G. and Sigman, R. (1982). The Use of Landsat for County Estimates of Crop Areas. Evaluation of the Huddleston-Ray and Battese-Fuller Estimators. Statistical Research Service Staff report no. AGES 820909, United States Department of Agriculture, Statistical Research Service.

Winings, S.B. (1982). Landsat Image Availability for Crop Area Estimation. Paper presented at the Eighth International Machine Processing of Remotely Sensed Data Symposium for Applications of Remote Sensing, West Lafayette, Indiana.

Zuttermeister, J.P. (1985). Evaluating TM Data for Statistical Research Service Crop Acreage and Production Estimates. Staff report no. RSB-85-02, United States Department of Agriculture, Statistical Research Service.

Received November 1988
Revised June 1990